A Publication of the American Psychological Association

The Journal of Abnormal and Social Psychology

Founded by Morton Prince, 1906

CONTENTS

ARTICLES

Cognition Without Awareness: Subliminal Influences Upon Conscious Thought 255
George S. Klein, Donald P. Spence, Robert R. Holt, and Susannah Gourevitch

A Study of Anxiety and Learning in Children 267
Richard R. Waite, Seymour B. Sarason, Frederick F. Lighthall, and Kenneth S. Davidson

The Effects of Different Social Feedback Conditions Upon Performance in Dyadic Teams 271
Seymour Rosenberg and Robert L. Hall

A Comparison of Interview and Observation Measures of Mother Behavior 278
Henrietta T. Smith

Neural Satiation, Reactive Inhibition, and Introversion-Extroversion 283
Allen Rechtschaffen

Body Image and Asymmetry of Body Reactivity 292
Seymour Fisher

Continued on inside front cover

Volume 57 NOVEMBER 1958 Number 3

CONTENTS—*Continued*

Time Orientation and Interpersonal Relations of Emotionally Disturbed and Normal Children 299
Anthony Davids and Anita Negrin Parenti

The Interaction of Achievement Motivation and Ability in Problem-Solving Success 306
Elizabeth G. French

Failure and the Subsequent Performance of Schizophrenics 310
Gordon W. Olson

The Effects of Plausibility on the Learning of Controversial Statements 315
Edward E. Jones and Riha Kohler

Vicarious Experience and Partial Reinforcement 321
Donald J. Lewis and Carl P. Duncan

The Expression and Inhibition of Hostility in Imaginative and Overt Behavior 327
Jack E. Hokanson and Jesse E. Gordon

Size Constancy as a Function of Personal Adjustment and Disposition 334
Neil Jenkin

Value Differences Among South African Students 339
Kurt Danziger

Family Patterns Correlated With Adolescent Personality Structure 347
Robert F. Peck

CASE REPORT

Some Dynamic Aspects of Posthypnotic Compliance 351
Milton J. Rosenberg and Charles W. Gardner

CRITIQUE AND NOTES

A Test of "Strain Toward Symmetry" Theories 367
Harry A. Burdick and Alan J. Burnes

Marital Satisfaction and Husbands' and Wives' Perception of Similarity in Their Preferred Frequency of Coitus 370
Paul Wallin and Alexander Clark

A "Pure" Measure of Perceptual Defense Uncontaminated by Response Suppression 373
Anne Mathews and Michael Wertheimer

Perceived Similarity and the Projection of Similarity: The Influence of Valence 376
Anthony J. Smith

Future Time Perspective, Optimism, and Academic Achievement 379
John E. Teahan

Hysterics and Dysthymics as Criterion Groups for the Measure of Introversion-Extraversion: A Rejoinder to Eysenck's Reply 381
John J. Sigal, Kelman H. Star, and Cyril M. Franks

Volume 57 NOVEMBER, 1958 Number 3

The Journal of Abnormal and Social Psychology

Editor: M. BREWSTER SMITH, *New York University*

Associate Editor: E. J. SHOBEN, JR., *Teachers College, Columbia University*

CONSULTING EDITORS

GORDON W. ALLPORT
Harvard University

URIE BRONFENBRENNER
Cornell University

ROGER BROWN
Massachusetts Institute of Technology

DONALD T. CAMPBELL
Northwestern University

LAUNOR F. CARTER
System Development Corporation

RICHARD CHRISTIE
Columbia University

MORTON DEUTSCH
Bell Telephone Laboratories

I. E. FARBER
State University of Iowa

LEON FESTINGER
Stanford University

DONALD W. FISKE
University of Chicago

ROBERT R. HOLT
New York University

IRVING L. JANIS
Yale University

HAROLD E. JONES
University of California

HAROLD H. KELLEY
University of Minnesota

LEONARD S. KOGAN
Community Service Society of New York

DANIEL J. LEVINSON
Harvard Medical School

GARDNER LINDZEY
University of Minnesota

DAVID C. MCCLELLAND
Harvard University

LEO POSTMAN
University of California

ELIOT H. RODNICK
Duke University

NEVITT SANFORD
University of California

SEYMOUR SARASON
Yale University

RICHARD L. SOLOMON
Harvard University

RENATO TAGIURI
Harvard University

Managing Editor: ARTHUR C. HOFFMAN

Promotion Manager: HELEN ORR
Editorial Assistant: BARBARA B. CUMMINGS

PUBLISHED BY

THE AMERICAN PSYCHOLOGICAL ASSOCIATION, Inc.

The JOURNAL OF ABNORMAL AND SOCIAL PSYCHOLOGY is devoted to both abnormal and social psychology. Its emphasis is on basic research and theory rather than the techniques and arts of practice. Abnormal psychology is broadly defined to include papers contributing to fundamental knowledge of the pathology, dynamics, and development of personality or individual behavior, including deterioration with age and disease. Articles concerned with psychodiagnostic techniques are evaluated with respect to their contribution to an understanding of the psychological principles of diagnosis, those concerned with psychotherapy are judged in terms of their contribution to an understanding of the therapeutic process. Case reports are considered in terms of their heuristic value. Reports which serve to pose or to clarify important theoretical problems, or which promise to be useful in teaching, are often published. From the social area this JOURNAL gives preference to papers contributing to basic knowledge of interpersonal relations, and of group influences on the pathology, dynamics, and development of individual behavior.

The JOURNAL is issued bimonthly in January, March, May, July, September, and November, two volumes per year. The two volumes average 784 pages. The subscription price per year is $16.00; foreign, $16.50. Single issues are $3.00.

Original manuscripts should be sent to the Editor:

DR. M. BREWSTER SMITH
Graduate Department of Psychology
New York University
Washington Square
New York 3, N. Y.

Fifty free offprints are supplied to contributors. Authors of prior publication articles receive no gratis offprints.

Subscriptions, address changes, and business communications should be sent to:

THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.
1333 Sixteenth Street N.W.
Washington 6, D. C.

Changes of address must reach the Subscription Office by the 10th of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee second-class forwarding postage. Other claims for undelivered copies must be made within four months of publication.

Printed and mailed by:

WAVERLY PRESS, INC.
Baltimore 2, Maryland

PUBLISHED BY
THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.
Mt. Royal and Guilford Aves., Baltimore 2, Md.
and 1333 Sixteenth Street N.W., Washington 6, D.C.

Entered as second-class matter February 15, 1954 at the Post Office at Baltimore, Maryland under the Act of February 28, 1925. Acceptance for mailing at special rate of postage provided for in Section 34.40 (e), P. L. and R. Additional entry at Washington, D.C.

Copyright 1958 by the American Psychological Association, Inc.

COGNITION WITHOUT AWARENESS: SUBLIMINAL INFLUENCES UPON CONSCIOUS THOUGHT[1]

GEORGE S. KLEIN, DONALD P. SPENCE, ROBERT R. HOLT, AND SUSANNAH GOUREVITCH

Research Center for Mental Health, New York University

EVIDENCES of cognitive activity without awareness have often been reported, both in the clinic and in experiment. Systematic interest in this phenomenon was first aroused by clinical investigations of dissociated states and posthypnotic suggestibility (2, 20, 21). From the beginning, a cardinal proposition of psychoanalysis has been that the greater part of mental life goes on outside awareness. In his detailed discussion of the way dreams are constructed unconsciously from fragments of recent experiences—usually indifferent ones, Freud (10) cited examples in which the "day residue" had been a fleeting glimpse of something that had hardly entered consciousness. Poetzl (19) first demonstrated experimentally that the manifest content of dreams could reproduce stimuli that had never been conscious. After showing his subjects (*S*s) landscapes tachistoscopically for about 1/100 of a second, he had them describe and draw what they had seen. What was *not* reported showed up the next day in *S*s' accounts and drawings of their dreams. Poetzl's investigations have been repeated in recent years by Fisher (8, 9) and by Luborsky and Shevrin (22) with substantially similar results.

Another line of experimentation in the last twenty years has concentrated on demonstrating that object discrimination is not contingent upon awareness (1, 11, 14, 15, 17). In more recent studies, interest in nonconscious intake of stimulus properties has centered around the special question of defensive *reactions* to subliminal registrations, as in alleged instances of "perceptual defense" (3, 6, 15, 16). In the main, however, experimental studies have not carried inquiry about nonconscious registrations to a next logical step: how and in what ways do subliminal registrations affect conscious thought and within what limits?

The issue of "subliminal reception" and thought is highlighted by another suggestion that may be derived from Poetzl's results: that incidental, peripheral, or subliminal registrations may have unique effects upon thought; that absence of awareness may actually promote certain forms of participation of percepts in ongoing thought. Poetzl believed that subliminal registrations are more likely to serve as vehicles for dream thoughts than are conscious percepts.

A serious barrier to studies of these issues has been one of method. In the technique of which we have heard the most in recent years, used in studies of "perceptual defense," *S* is confronted with supposedly critical stimuli which are gradually raised from a low level of articulation to clearly recognizable levels. In this method, the nonconscious registration and the defensive reaction to it are *both* inferred from deviant recognition times. The threshold technique of these studies provides little if any information concerning the processes that determine the threshold measure. A new technique seemed necessary before nonconscious perceptual and cognitive processes could be studied experimentally in more detail.

To investigate these problems, we developed a method that records the impact of a subliminal registration upon the nature, extent, and quality of an *S*'s conscious impressions of a

[1] An abridged account of this study was presented at the Sixty-Third Annual Convention of the American Psychological Association, 1955.

The study is one of a series supported by a grant from the National Institute of Mental Health (Project Code Number: M-782, Cognitive Processes Below the Level of Awareness), and was carried out in the Psychosomatic Laboratory and Department of Psychiatry, School of Medicine, State University of New York. We are grateful to Robert M. Dickes, who was then Director of the Psychosomatic Laboratory, for generously making available the Laboratory's facilities and giving us access to the student body of the medical school for experimental *S*s, and to Michael Tobin and others of the Laboratory staff for their freely offered assistance at various points in the study. Our appreciation also goes to Jan Frank of the Department of Psychiatry for useful exchanges of views regarding interpretations of the data.

We wish to thank Sheldon Bach, Reeva Safrin, and Joan Havel who rated the protocols; and Bessie Boris, Worden Day, and Leo Garel, the artists who drew the stimuli and helped develop the check list for evaluating the *S*s' drawings. We are also indebted to Harriet Linton and Gudmund Smith for their assistance as judges.

perceived stimulus. The technique does not require *S* to report directly on the critical subliminal stimulus; rather, the influence of subliminally presented figures is determined by the study of associated changes in *S*'s responses to a consciously seen figure.

The technique is based upon a phenomenon described by Werner (25) and later by Cheatham (5). Expose in quick succession a pair of figures at the same point on the screen of a mirror tachistoscope. If the two follow each other in sufficiently rapid succession and if the second appears in a slightly longer exposure, only the second of the pair (*B-figure*) is consciously seen. Complete "obliteration" of the first figure (*A-figure*) occurs, even when it is exposed so long that it could be clearly recognized if it were not succeeded by the second figure.

Werner and Cheatham considered relative contour strength and spatial disparity of the critical A- and B-figures to be critical variables in accounting for the phenomenon. But if the phenomenon can be obtained with complex figures when such congruities of contour do not exist, the way is opened to use it as a means of studying issues of "subliminal registration." Pilot studies strongly suggested that the A-figure, even though not seen, exerted an effect upon conceptions of the B-figure. If so, it would be possible to vary systematically the *meaning* of the masked A-figure and to look for distinctive effects of such meanings upon impressions of the consciously perceived second figure.

The present study centered upon two main questions: (*a*) Will impressions of a pictured person (B-figure) of ambiguous sex be significantly affected by subliminally presented sexual drawings? (*b*) Will impressions of the B-figure be differently affected by realistic and by symbolic sexual pictures shown as A-figures? When sexual connotations of a form are presented symbolically and thus disguised, its effects upon thought may be distinctively different from those of uncamouflaged, realistic forms.

METHOD

Apparatus

An electronically controlled mirror (Dodge) tachistoscope was used which consists of three fields: a blank field on which *S* is ordinarily focussed, and two stimulus fields.[2] The rest of the interior is painted flat black or covered with black velvet. The blank field consists of a white cardboard on which is drawn a black rectangle circumscribing the area where the stimulus figures appear.

Each stimulus compartment is equipped with two mercury argon tubes surfaced with a special phosphor coating that produces a daylight white color, with rapid onset and quick decay so that the light output closely approximates a square wave form even at very brief exposures.[3]

The illumination of each field can be independently varied. All fields were adjusted to equivalent brightness by means of a photometer placed at the viewing aperture. In the present study, the brightness values of the three fields were fixed at .8 apparent foot candles for the range of exposure times used (as determined by a Macbeth Illuminometer).

The electronic timing unit fires the light source in the three fields independently in predetermined sequences. Durations of the two stimulus fields can be varied continuously from .002 to 5.9 seconds. The accuracy of the time intervals is ± 1 per cent above .10 sec. and ± 2 per cent below .10 sec. The two stimulus fields may be exposed singly or in succession with no intervening intervals. The blank field reappears after each stimulus or stimulus sequence. The start of the sequence is controlled manually.

Each stimulus compartment accommodates 7″ square stimulus cards. The stimulus fields are arranged in the box at a 90° angle from each other at an optical distance of 3.5′ from the eyepiece. Coated low-reflection mirrors with a transmission-reflection ratio of approximately 50-50 are carefully fixed in the box so as to transmit the two stimuli over a common optical path to *S*'s position. The placement is such as to equate the retinal sizes of each of the stimulus fields.[4]

Design and Procedure

The B-stimuli, shown in Fig. 1, were line drawings of three human figures whose sexual identity was intended to be uncertain. The A-stimuli were India ink line drawings of male genitals; female genitals; a "masculine symbol" (trylon and perisphere); and a "feminine symbol" (a shield-like triangular form). The A-figures are shown in Fig. 2. The A-stimulus, whether realistic (male or female genitals) or symbolic, was always placed in the center of the field. Thus, the genitals were anatomically displaced with respect to the B-figure.

The symbols were chosen from among a larger number of possibilities after we found that pretest *S*s,

[2] A complete description of the tachistoscope and electronic timing units is obtainable from Fred and Charles Snyder, The Menninger Foundation, Topeka, Kansas, who designed and constructed them.

[3] We are grateful to Lawrence Karlin for his help in finding white-light tubes with desirable onset and decay attributes and for checking their wave forms. The tubes are described in Karlin (12).

[4] The adjustable card carriers make it possible to displace each of the two fields laterally when this is desired. The maximum lateral displacement of the two fields is one-half of a 7″ card.

FIG. 1. B-STIMULI

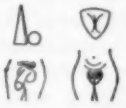

FIG. 2. A-STIMULI

asked to associate to them, gave consistently female sexual associations to the female sexual symbol and consistently male ones to the trylon and perisphere (though for the latter, obvious sexual meanings appeared only later in the flow of associations).

The four A-figures were paired with the three B-figures and shown subliminally to different *S*s and in different pairings in a balanced design shown in Table 1. A control condition was also included in which B-figures were shown without a preceding A-figure.

Each *S* was tested in two sessions. In Session I he was shown a B-figure with no preceding A-figure (control condition) and then another B-figure with a preceding subliminal A-figure (e.g., male genitals or female genitals; male symbol or female symbol). For half the *S*s the control-subliminal sequence was reversed. In Session II, which followed the first by three to ten days, *S* saw a mirror image[5] of the *same* B-figure with the opposite member of the first subliminal series, and then the third B-figure paired with a supraliminal A-figure.[6] The balanced design provided for equal distribution of all pairings and conditions among the group of 24 subjects. The conditions and numbers of subjects in each are summarized in Table 1. *S*s were assigned to conditions randomly. In the case of 17 *S*s it was possible to repeat the control (blank) condition in the second session as well. Following the test conditions in the second session, ascending recognition thresholds were determined for the A-figures presented *alone* with each *S* tested on stimuli that had been shown him as A-figures in the subliminal conditions.

B-figures were always exposed for .33 second, far above recognition threshold and providing *S* with a clear but brief look at the stimulus. The A-figures were exposed in the subliminal condition at much briefer exposures, varying slightly from *S* to *S* within a range of .02 to .035 second. The exposure time for subliminal A-figures was tailored to each *S*'s recognition threshold for geometric figures (a diamond and a cube) which was determined prior to the critical experimental trials in the first experimental session. To make allowances for discrepancies of form between the geometric figures and the critical A-figures, a constant, arbitrary value of .015 second was added to the higher of the *S*'s recognition threshold for the geometric figures. For instance, if *S*'s recognition threshold for the diamond was .005 second and for the cube .01 second, addition of the constant gave a value of .025 second, which was then fixed as his exposure time for all A-figures shown to him in the subliminal conditions. In all A-B pairings, the B-figure immediately succeeded the A-figure so that the only phenomenal change was the sudden brief appearance of the dark lines of the B-figure in an otherwise constant field. Between the exposure of stimulus pairs, the blank field remained on steadily. General illumination in the experimental room was low.

The first experimental session began with the measurement of *S*'s ascending recognition threshold for the

TABLE 1

EXPERIMENTAL DESIGN

(N = 4 for each of six conditions)

Session 1		Session 2	
Subliminal[a]	Control	Subliminal	Supraliminal[b]
M gen-I	II alone	F gen-I	M/F sym-III
M gen-II	III alone	F gen-II	M/F sym-I
M gen-III	I alone	F gen-III	M/F sym-II
M sym-I	II alone	F sym-I	M/F gen-III
M sym-II	III alone	F sym-II	M/F gen-I
M sym-III	I alone	F sym-III	M/F gen-II

M gen = male genital A-stimulus
F gen = female genital A-stimulus
M sym = male symbolic A-stimulus
F sym = female symbolic A-stimulus
M/F = male or female
I = B-picture (face)
II = B-picture II (figure with arms folded)
III = B-picture III (figure with back turned)

[5] The repeated B-figure was shown in mirror-image in order to make any perceptual differences caused by the change of the A-figure more plausible to the *S*. Care was taken to balance all variables so that there was no confounding of sequence (M to F or F to M), type of A-figure, and orientation of B-figure. Since no systematic differences in the ratings of drawings of the B-figures were associated with their being in the original or the mirrored orientation, no further reference will be made to this factor.

[6] Results with the supraliminal pairing of A- and B-figures are not included in the present report because of an unanticipated flaw in the design. It turned out that reactions to the different B-figures differed so much that all comparisons we could make between effects of subliminal and supraliminal pairings were confounded with B-figure differences. Separate analyses for the different B-figures were precluded by the small number of *S*s in each stimulus in the supraliminal condition. We plan to compare supraliminal and subliminal effects more intensively in another study.

[a] Half the *S*s had a male A-stimulus in Session 1, half a female A-stimulus. When male was presented in Session 1, female was presented in Session 2 and vice versa. Also, half the *S*s in each sequence had the subliminal presentation first in this session, half had the control first.

[b] The A-stimulus presented to seven *S*s deviated from the general design: they were given one of the A-stimuli which they had been shown earlier under subliminal conditions.

cube and the diamond. *S* was comfortably seated so that he could rest his forehead against the eyepiece and look with both eyes (keeping his glasses on if he wore them) through the black tunnel of the tachistoscope to the moderately bright, blank fixation field. After the subliminal A-stimulus exposure time was found in the manner described above, *S* was next told:

> You are going to see a figure that will appear briefly. I will show it 10 times and each time I will say "Ready," just before I show it to you. I want you to look at it carefully because later on I will ask you to describe the personality of the figure, the type of person the figure seems to be. I'll also ask you to draw it.

There were then 11 exposures (one to prepare *S* plus 10 more) at 2-sec. intervals of either a subliminal A-figure followed by a B-figure, or a B-figure alone (control condition). *S* then went to an anteroom where he filled out rating forms describing the "personality" of the B-figure, made a free drawing of it, and was interviewed on other aspects of his impression.

In the second session, *S* was shown a new A-figure in a subliminal pairing with the same instructions. He was also shown a new pair of A- and B-figures in a supraliminal presentation.

To review briefly the test conditions for an individual *S*: *Session I:* recognition thresholds were determined for geometric figures. Then followed the subliminal A-B pairing and the control B-figure (or vice versa). *Session II:* subliminal A-B pairing, followed by the supraliminal A-B pairing and the control B-figure again. Finally, recognition thresholds were obtained for each of the A-figures that had been shown to *S* in the subliminal condition.

Subjects

*S*s were 24 male medical students in the third year of training. During the period of the experiment, they were engaged in a one-month psychiatric "clerkship" and their participation was made part of the course requirement. They were told only that it was an experiment in perception.

RESULTS

Principal Response Variables

Our major hypothesis was that impressions of the B-figures would incorporate expressive qualities and meanings of the subliminal A-figures. Our first, simple hypothesis was that the subliminal figure would cause *S* to see the B-figure as having the corresponding sex. Since all *S*s tended to see all B-figures as males, we had to look for effects in subtler measures than an indication of sex. Thus, we expected that a B-figure, presented contiguously with the male genital, would acquire more attributes that were distinctively associated with the male genital than it would when paired with the female genital. Essentially the same trends were expected with the symbols.

The *S* recorded his impression of the B-figures in two ways: verbal descriptions (a personality check list) and drawings of the B-figures.

Personality check list. This check list consisted of 69 pairs of adjectives, e.g., "modest-boastful," each pair forming the extreme positions of a 4-point scale. Items were chosen that seemed applicable to all three B-figures and at the same time might be sensitive to the effects of the subliminal A-figure. Osgood's list (18) was a principal source of items.

The same personality check list was employed in developing scoring keys on the expressive qualities of the A-stimuli. Twelve male judges, graduate students in clinical psychology, described all four A-figures. We pooled the description of each panel, disregarding differences between the "somewhat" and "very" ratings, and compiled into clusters those items that at least 10 out of 12 judges applied to the male genital, the female genital, the male symbol, and female symbol. The clusters were further paired by eliminating items common to each pair, with the result that the final clusters represented distinctive, nonoverlapping attributes which the given stimulus provoked in the majority judgment. The four A-figure keys were used in the analysis of *S*s' descriptions of B-figures on the check list and are presented in Table 2.

Drawings. In the subliminal conditions, each *S* made two drawings of the B-figure, one in the male A-figure condition, one in the female A-figure condition. Each pair of drawings was compared in *manliness* by five psychologists who made a "more than"–"less than" judgment. The number of judges making the discrimination in a particular pair of drawings was used to weight the difference; the more judges agreeing on a pair, the higher the relative masculinity score. Thus, highest rank went to an *S* whose drawing in the male genital condition was judged more manly by all five judges. Lower ranks were given to *S*s whose drawings in this condition were judged more manly than the other by less than unanimous decision (i.e., 4-0, 4-1, 3-0, 3-1, 2-0, 2-1). Male genital drawings judged *less* manly were ranked in reverse order in the same fashion. In this way, the 12 pairs of drawings could be ranked for comparative masculinity in each condition. (We will refer to this measure

TABLE 2
A-STIMULUS KEYS
(Unique characteristics of the A-figures as judged by 10 out of 12 judges)

Male Genital	Female Genital	Male Symbol	Female Symbol
boastful	womanly	brave	womanly
brave	peaceful	clear-thinking	loving
happy	loving	manly	subtle
manly	contented	strong	warm
strong	delicate	ambitious	kind
ambitious	soft (opposite loud)	decisive	rounded
decisive	intuitive	active	generous
active	smooth	rugged	
rugged	sweet	clean	
outgoing	soft (opposite hard)	calm	
dominant	vivacious	outgoing	
hard	kind	dominant	
large	rounded	vigorous	
optimistic	generous	independent	
mature		hard	
		large	
		effective	
		angular	
		mature	

as Comparative Masculinity.) This method of evaluating the drawings assumed that the judges' conception of manliness would be related to the effects of the A-stimuli. A drawback of this score was that it was based on a comparison of *S*'s own drawings only. One cannot tell how *S*s compared with each other in respect to the manliness of their drawings.

Therefore, another technique of assessing manliness was used. Drawings of B-figures made by *all* *S*s in each subliminal condition were ranked in respect to manliness. (We will refer to the measure as Relative Masculinity.) Two judges made each ranking and their combined ranks was the final score for each *S* on a given condition.

Subliminal Effects: Realistic A-Stimuli

Before we go into the effects of the subliminal stimuli, consider for a moment *S*'s task: he was given a short series of glimpses of a line drawing; a few minutes later he had to make a rather detailed set of ratings and then draw his fast-fading impression of what he had seen. We expected that some attributes of the A-figure would be incorporated in *S*'s check-list description of the B-figure. Therefore, when a B-figure was paired with the male genital, there should be more items from the male genital key in descriptions of the B-figure then when the same B-figure is paired with the female genital.[7] An agreement is scored if *S*'s reaction to the B-figure on any item agrees with the key. The degree to which the male stimulus was *uniquely* effective in provoking the qualities associated with it in the key was found by subtracting agreements with the male key in the female stimulus condition from the number of agreements in the male condition; these differences were ranked for the 12 *S*s.

Agreements with the A-figure key were higher in the male A-stimulus condition for six *S*s, lower for three *S*s, and unchanged for three *S*s. Assuming that the key is a reasonable summary of connotations provoked by the A-stimulus, it is evident that not all *S*s incorporated these qualities in their descriptions of the B-figure. *S*s were *consistent*, however, in their tendency to do so or not. It is interesting that all the reversals of this tendency were consistent, i.e., if *S*s responded to the male A-figure with qualities of the female genitals, they responded to the female A-figure with qualities of the male genital. Hence, it is not surprising that the degree to which *S*s brought qualities of the stimulus key into their descriptions of B-figures is correlated .82 in the male and female genital A-stimulus conditions. If the male genital induced *S* to incorporate male genital qualities, the female genital was likely to be equally effective in coloring impressions of the B-figure with its associative qualities.

Further evidence for a consistent reaction of the *S*s in the two subliminal presentations of the genital stimuli comes from the drawings. As the correlations in Table 3 attest, *S*s who attributed characteristics of the subliminal male genital A-figure to the paired B-figure also produced the most manly drawings in response to the male A-stimulus (i.e., they ranked highest on Comparative Masculinity). This appropriate differentiation of drawings according to the A-figure exposed also correlated significantly with the check-list measure of the effectiveness of the female genital. Thus, *S*s differed considerably in their sensitivity to influence by the subliminal genital stimulus,

[7] Comparisons of the subliminal conditions with the blank condition were difficult to make because the B-figure paired with the blank was not the one presented to *S* in the subliminal conditions. Since different B-figures provoked varying impressions on the check list, it was difficult to assess the unique effects of subliminal stimuli in such a condition.

TABLE 3
RANK-ORDER CORRELATIONS BETWEEN MAJOR VARIABLES

Realistic Stimuli	
Male genital effect—Check list vs. Female genital effect—Check list	.82***
Male genital effect—Check list vs. Comparative masculinity—Drawing	.64**
Female genital effect—Check list vs. Comparative femininity—Drawing	.54*
Male genital effect—Check list vs. Relative masculinity—Male genital drawing	.51*
Female genital effect—Check list vs. Relative femininity—female genital drawing	.18
Symbolic Stimuli	
Male symbol effect—Check list vs. Female symbol effect—Check list	.34
Male symbol effect—Check list vs. Comparative masculinity—Drawing	−.12
Female symbol effect—Check list vs. Comparative femininity—Drawing	−.64**
Male symbol effect—Check list vs. Relative masculinity—Male symbol drawing	.24
Female symbol effect—Check list vs. Relative femininity—Female symbol drawing	−.27

* $p < .10$.
** $p < .05$.
*** $p < .02$.

Note.—For 12 cases, a rho of .51 is significant at the .10 level, a rho of .57 is significant at the .05 level, and a rho of .71 is significant at the .02 level, all two-tailed tests.

but they tended to be consistent in the amount of effect they showed in both graphic and judgmental behavior (drawings and check-list ratings).

We can also look for consistencies of reaction within a single session rather than in comparisons of sessions, as in the data just cited. The Relative Manliness of the drawing was correlated with the check-list measure of the effectiveness of the male genital A-figure at a near-significant level, though this same consistency did not show up in the session when the female genital was the A-figure (Table 3). Evidently the manliness induced by the subliminal male genital is the effective agent in producing the differences between the two *S*s' drawings with respect to masculinity. Because the B-figure is not truly ambiguous but usually seen as a man (so described by eight out of twelve *S*s in the blank condition) it may be more susceptible to the effect of a male A-stimulus than to a feminine A-stimulus. Of course, the male genital may be generally a more powerful subliminal stimulus for this sample of subjects; however, this seems unlikely in view of the high correlation (.82) between sensitivity scores on the check list under the two conditions. At any rate, the subliminal female genital did not produce the expected increase in the femininity of the B-figure *drawing*.

Self-image. One source of variance is the masculinity of *S*'s self-image. Several weeks after the experiment, *S*s were asked to describe themselves on the same personality check list used in the main experiment. Using the two genital keys as extremes of masculinity and femininity, we tallied agreements with *S*'s ratings of self on the same check list. *S*s with relatively *feminine* self-ratings made very masculine drawings of the B-figure following the subliminal female genital. Relative Masculinity of drawings in this condition correlated .62 with relative femininity of self-image ($p < .10$, two-tailed test. All significance levels below are two-tailed). Conversely, a relatively masculine self-image is associated with femininity of the drawings (same correlation). In other words, the female genital provoked the expected feminine drawings in those *S*s whose masculine self-picture was fairly clear, whereas it produced a more masculine drawing among *S*s with a relatively large feminine component in their self-conception.

Subliminal Effects: Symbolic A-stimuli

In order to compare effects of the two main types of A-figures, analysis of the symbols followed the same lines described for the genitals, and starts with the two stimulus keys. As Table 2 shows, the male symbols represented largely masculine qualities to our panel of judges, supporting our assumption that the trylon and perisphere is also a male symbol. To a much lesser degree, the so-called feminine symbol was assigned feminine connotations. The relatively small number of items that *uniquely* described the female symbol key indicates the low degree of uniformity of meanings possessed by this abstract form.

Neither A-figure had a distinctive effect consistent across *S*s on check-list impressions of the B-figure; in this respect, reactions to symbols resembled those to realistic figures. Six *S*s incorporated more aspects of the subliminal male symbol in their B-descriptions under the male symbol condition; four incorporated fewer aspects, and two showed no change. Only four *S*s showed the expected effect in the female symbol condition. Also indicative of this inconsistency between stimu-

lus reactions is that reversals of expectation in one condition did not always occur with reversals in the other condition. This contrasts with the high inter-*S* consistency on the genital stimuli. The correlation between the male and female symbols with respect to the distinctive effectiveness of the A-stimuli was not significant—only .34.

The drawings were more in accord with the hypothesis of a direct subliminal influence of the symbols. In nine of the twelve *S*s, drawings of the B-figure in the male symbol condition were more masculine than corresponding drawings under the female condition ($p < .04$, sign test two-tailed). The male symbol, sharing many qualities with the male genital according to the panel of 12 judges, appeared to influence the B-figure directly and to provoke a relatively more masculine drawing. Moreover, the effect was much more uniform than that of the male genital which produced more masculine drawings in only 7 out of 12 *S*s.

Comparative sexuality of the drawing tends to be negatively rather than positively related to responsiveness to symbol attributes on the check list. Comparative Masculinity of the B-figure drawing following the male symbol is correlated −.12 with agreement with the male symbol key; Comparative Femininity of the B-figure drawing following the female symbol is correlated −.64 with agreements with the female symbol key. Evidently, the subliminal female symbol provoked a relatively more feminine drawing of the B-figure in *S*s who described it as more masculine on the check list, suggesting that the effect of the subliminal stimulus was expressed in only one of these response channels for most *S*s.

Possibly the differences between results with check list and with drawings stem (*a*) from the distinctive linear qualities of the two symbols and (*b*) from the differences in covert and overt meanings of the two symbols. A distinctive feature of the male symbol is its long triangular form; a distinctive feature of the female symbol is its curving lines and general symmetry. The genital stimuli, by contrast, have many more figured similarities. These differences in linear quality of the symbols may have been especially evocative of different graphic responses, in ways that were relatively independent of the sexual meanings. On the other hand, the divergent meanings that *S*s assigned to the female symbol make for a less uniform, relatively unstereotyped (verbal) impression (see Table 2). The male symbol clearly had an obvious, social meaning (trylon and perisphere) which is certainly more accessible than its covert sexual connotations. The female symbol, not having such a clear socially stereotyped meaning, produced a considerable variety of interpretations and was hence susceptible to idiosyncratic interpretation.

The self-image data suggest that certain *S*s are more sensitive to the feminine aspects of the female symbol, namely *S*s with a highly masculine self-image. *S*s whose self-ratings contained a large number of items from the male genital key were highly sensitive to the female symbol; many of the distinctive attributes of the female symbol appeared in their descriptions of the associated B-figure. Correlation between masculinity of self-image and sensitivity to the female symbol is .75, $p < .02$. Evidently for certain *S*s, a clear sexual meaning in the female symbol is subliminally effective. However, for the symbol group as a whole, form seems to be a more clearly effective determinant of the response and tends to obscure the particular contribution of the sexual meaning.

Recognition Thresholds

We will now consider whether or not indices of threshold sensitivity to the A-figures were correlated with the behavioral effects that were central to the present study. Thresholds for both sets of A-figures were obtained at the end of the second experimental session, with threshold defined as that exposure level at which *S* could accurately identify the A-figure. The lowest exposure was 4 milliseconds and exposures were increased in an ascending series in logarithmic progression until recognition was achieved. On the genitals, we considered recognition to be complete when *S* said "penis" or "male organs" or "vagina" or "female organs"; on the symbols, when *S* was able to say "trylon and perisphere" or "circle and triangle" for the male symbol, and accurately describe the inner details of the female symbol. Order of presentation was random for each group of 12 *S*s, with 6 *S*s receiving the male symbol first and 6 the female first; 5 the male genital first and 7 the female genital first. For

each *S*, thresholds were obtained only for the subliminal A-figures used in the main experiment.

We will consider first the relationships between the genital thresholds and check-list reactions to the subliminal male and female pairings. Threshold sensitivity was unrelated in any significant way to distinctive responsiveness to the subliminal stimulus. Sensitivity was, however, related to the degree to which *S*s responded differentially to the A-figures on the check list, regardless of whether their reactions conformed to the stereotype of the A-figure. Average threshold for the two genital A-figures was significantly correlated with this measure of differential responsiveness on the check list (rho = $-.61$, $p < .05$), with *S*s having high thresholds for both genital figures giving the least differentiated responses to the two genitals in the subliminal condition. Sensitivity to the male genital seems to be the major factor in this relationship, for male genital thresholds were significantly correlated with differential responsiveness (rho = $-.59$, $p < .05$) while female genital thresholds were not (rho = $-.24$).

Turning to the A-figure effects on the drawings, we find a more complex relationship with thresholds. High threshold alone is not related to graphic measures of subliminal sensitivity, but *S*s with either extremely high or extremely low thresholds to both genital stimuli drew noticeably more masculine figures in the male subliminal condition than did *S*s with more average thresholds ($p = .05$, Mann-Whitney test). The more deviant the threshold, the more the drawing reflects the subliminal figure.

Recognition thresholds for the symbols were related in a markedly different manner to subliminal effect scores. Whereas genital thresholds showed some relationship to check-list responses, threshold level for the symbols was related only to the drawings. Low threshold for the male symbol was correlated .52 ($p < .10$) with Relative Femininity of the drawing in the female symbol condition, and low threshold for the female symbol correlated a .60 ($p < .05$) with Relative Masculinity of the drawing in the male symbol condition. The direction of these relationships is not readily interpretable. We are able only to say that high threshold for one symbolic figure is related to sensitivity to the opposite symbolic A-figure in the subliminal condition.

Two general trends stand out. Recognition thresholds for the genitals and the symbols are each related to subliminal behavioral effects, but in different ways. Check-list responsiveness is generally related to thresholds for the genital figures while responsiveness on the drawings is related to thresholds for the symbols. It may be that the threshold for the genital figures is primarily a threshold for meaning and for this reason relates most directly to the check-list indices of subliminal effect. Symbolic thresholds are more clearly thresholds for form, and understandably are related more directly to the masculine and feminine properties of the drawings which were judged on a largely formal basis, using such categories as roundedness of line, symmetry, etc.

DISCUSSION

The main implication of the findings is that stimuli may acquire psychological representation without *S*'s awareness, and subsequently affect cognitive and behavioral processes to a measurable degree. There is little question, however, that the principal effects appeared only in subtle measures and were complicated by idiosyncratic interactions among the variables. Does such a seemingly unstable phenomenon merit further concern? In our opinion it does. Important theoretical issues are perhaps concealed by the fact that the *unconscious* influences on perceptions *are* small under ordinary circumstances. Were it possible now to maximize the effects by manipulating appropriate variables, such issues might be brought into relief.

We will consider some of the variables and issues that were especially highlighted by the present findings, and indicate some possibly fruitful directions for further study.

1. *Meaning and form in the effects of A-figures.* In their simplest guise, the results suggest that the A-figure interacts in some manner with the B-figure to produce a response that combines or in some way uses the elements of the two stimuli. Both meaning and form seem to be crucial determinants; the different effects of the two variables are highlighted in the results with realistic and symbolic A-figures.

The most obvious difference in the effects of the two types of A-figures was that the genital stimuli produce consistent results on both drawings and check list, whereas the symbols' effect is pronounced only on the drawings. In the symbol condition, the one sizable trend was that a significant majority of *S*s drew the male symbol B-figure as more masculine than the female symbol B-figure. Group effects on the drawing (and also the check list) were nonsignificant with the genital A-stimuli. However, individual differences were more meaningful, as evidenced by significant correlations between drawing and check list, and indicated that the A-stimuli were effective though in idiosyncratic ways.

The greater consistency of results for the genital condition appears to be related to the greater coherence of the genital A-stimuli, in the sense that the form of the genitals is inextricably tied to their meaning. Considered simply as stimuli, they are unambiguous; there is no escaping their crude sexual message. If we assume that formal effects are represented primarily on the drawings, and content effects on the check list, the high correlation between the two measures seems highly reasonable.

The symbols are quite a different set of stimuli. The spherical triangle (see Fig. 2) does not unambiguously represent femininity; a glance at the group consensus shown in Table 2 points up the lack of agreement in how the judges interpreted the female symbol. There is less *invariance* between meaning and form in the case of the symbols; they have other connotations besides sexual ones, and we have no guarantee that sexual connotations were primary. The forms might have sexual equivalents for some *S*s and not for others, and for this reason, correlation between drawing and check list would be low.

Because the connotations of the symbols are diverse, the forms of these stimuli are more independent as stimulus properties. Now it is possible to assume that purely formal aspects of the A-figure interact with formal aspects of the B-figure so that the response represents the direct effect of one contour on another. The more masculine drawings following the male symbols may represent such an interaction; the size of the effect may stem from the fact that because form is the main variable, individual associations to the symbols are not involved.

If the more uniform drawing results do indeed reflect the formal properties of the symbols, it would seem that a stimulus with pronounced form and disguised or ambiguous content has a definite subliminal advantage. Perhaps ambiguity of meaning helps to bypass individual associations which tend to complicate the group results. It is also possible that formal stimulus properties are less affected by the high exposure speed used in the tachistoscopic presentation, and preserve more of their original stimulus properties, whereas more meaningful stimuli may be differently affected by different speeds.

2. *Nature of subliminal interaction.* How, exactly, do the A and B processes interact? At least two possibilities are suggested by the results. In some *S*s, elements from A and B stimuli fuse to produce a condensation, as when the drawing of the ambiguous B-figure becomes more masculine when it follows a male symbol. In other *S*s, the A and B elements are kept separate, with the properties of the A-figure serving as a ground against which the B-figure is judged. In this case, the ambiguous B-figure is seen as more feminine following the male genital, and its difference from the sex of the A-figure is emphasized. As noted above, some *S*s included attributes of the genital figures in their impressions of the B-figures, others excluded such attributes, and others were not affected, with a high consistency in individual responses. If *S* included attributes of the male genital, he did the same with the female genital. These consistencies strongly suggest that inclination to condense or contrast the A and B stimuli may be a personality variable. Different responses to the subliminal genitals seem to depend partly on the feminine component of the self-image, and more generally, on *S*'s overall sexual adjustment. We know that *S*s with a well-established masculine self-image tend to incorporate elements of the subliminal female genital stimulus in their drawings of the associated B-figure, whereas *S*s with less masculine self-image tend to exclude feminine elements. Strength of the masculine self-image in male *S*s is certainly related to general sexual adjustment; a low masculine self-image hints at specific conflicts in this area. These

conflicts may be aroused by a subliminal female genital and lead *S* to deny the female aspects of the B-figure. We suggest the possibility, therefore, that the "choice" of condensation or contrast-formation as the way a subliminal stimulus interacts with conscious cognition may vary with the stimulus material, but that for stimuli like ours, it may be a function of defenses associated with self-identity.

Interactions of the A and B processes may be represented in different qualitative experiences of the B-stimulus. The interaction may be reflected in *S*'s perception of the B-figure, in judgments, in feelings, or in some combination of these. It may occasion a variety of verbal and motor reactions. Thus, the interactions between the two processes need not be represented only perceptually but may appear in various forms of behavior or experience.

The qualitatively different outcomes of the A and B interaction are communicated to *E* through different channels, some of which involve awareness, others not. It seems worthwhile to distinguish expressive channels of report from the experiential and nonexperiential qualities of *reaction* to B. In the present study for instance, it cannot be said to what extent the check-list responses or the drawings represent perceptual experience, judgments, or imaginal elaborations. Moreover, some forms of report convey certain qualities of the experience more easily than other forms of report. Description of a face in a set of fixed categories offers very limited means for conveying a variety of reactions as compared to drawings. This differential sensitivity may explain the inconsistency between the drawings, check list, and other forms of report, and suggests that knowledge of subliminal processes waits on the development of more sensitive measures.

3. *Recognition thresholds and behavioral effects.* Comparison of the recognition threshold data of the present study with changes in the B-figures points up some of the drawbacks of studies which have sought evidence of "defense" solely in recognition threshold level (e.g., 6, 15, 16). Our findings suggest that the recognition threshold is only a limited and for certain stimuli even an insensitive indicator of the personal significance of the stimulus. For instance, the symbols had rather clear effects both in the check list and in the drawings, but their thresholds were not especially elevated or depressed, and individual differences in thresholds were related to differences on the drawings alone and then only in ambiguous fashion. Even the more widely varying recognition thresholds for the genital stimuli did not reliably indicate their potency in other areas of response. As a whole, the results suggest that the modes of cognitive elaboration which a stimulus may undergo below awareness are only coarsely represented in the narrow datum of recognition level.

4. *Discrimination and registration.* We have assumed that in the masking conditions of the present study, exposure of the A-stimulus occurred at limits beyond *S*s' capacity to discriminate the figures, either in part or in whole. Our findings, like those of Fisher (8, 9) and Luborsky and Shevrin (22) support a conception of *registration* (reception, organization of intake, trace activation) without partial or complete perception.

On the other hand, there are those who maintain that it is unnecessary to assume two processes (4, 7, 23). Their argument runs something like this: Only a discriminable stimulus can affect thought, and discrimination always depends upon at least some information being within *S*'s conscious grasp. Therefore, if the A-stimulus has exerted any effect it could only be because just enough of its properties were partly or dimly accessible to make possible the discrimination required for the stimulus to have any further behavioral impact. Adams, in a recent review of studies that purport to demonstrate discrimination without awareness, concluded that "... the only type of behavior without awareness which can be easily reproduced on the basis of published reports is the classical type, in which *S* knows *what* he is supposed to be discriminating, but does not know *that* he is discriminating, because of the absence of the usual sensory experiences to which he is accustomed under the given type of stimulation" (1, p. 402). In the present study, *S* certainly did not know *what* he was supposed to be discriminating in the "subliminal" stimulus, nor even *that* he was responding to a subliminal stimulus, for the test conditions did not require him to respond *directly* to it (in contrast to the usual discrimination-without-awareness studies) but

only to a contiguous, consciously seen form. In spite of this, we were able to observe changes in reported impressions of the consciously seen figure, and to relate these to qualities of the A-figures.

The crucial question that remains is whether *S was* partially aware of the A-stimulus but did not report it. The issue is clearly drawn. The assumption that conscious perceptibility is not a necessary condition for registration has considerable implications for a general theory of cognition. But before such an unusual assumption can be accepted, it is necessary to rule out the simpler explanation that *S* has some partial awareness of the so-called subliminal stimulus even under the masking conditions, and is simply responding to that.

Can we be sure that the A-figure was in fact subliminal, in the sense that the A-figure was not even partly discriminable? The present experiment did not explicitly aim to explore this possibility, and we must grant that our findings do not settle the matter. None of the *S*s gave any indication that they were aware of another form other than the B-figure, but of course our method of report did not completely rule out this possibility.[8] In another study, we propose to obtain discrimination thresholds under a variety of conditions designed to produce optimal circumstances for at least the partial recognition of a stimulus. Thresholds obtained in this way will then be used to set stimulus exposures at a level where we can safely presume that not even partial discrimination is possible. We shall then explore the behavioral effects of such subliminal exposures in a manner patterned after our present procedure as well as that of Fisher (8, 9) in his closely related work.

SUMMARY

A method is presented for studying cognitive effects of stimuli that do not give rise to conscious percepts. The *S* is shown a drawing of a person ambiguous as to sex, in a mirror tachistoscope, for .33 second. This exposure is preceded by a briefly exposed stimulus (approximately .027 second) which is not consciously perceived. The first, subliminal presentation is called *A-figure*, the second one *B- figure*. In this experiment, the B-figure was always a sexually ambiguous drawing of a person.

In a balanced design using 24 male medical students, this technique was used to study the effects of four A-figures—a male symbol, a female symbol, male genitals, and female genitals—on three different B-figures. In the first session, *S* was shown a B-figure preceded by a subliminal A-figure; he then drew the B-figure and rated it on 69 four-point adjective scales. Similar ratings and drawings were obtained after each exposure of a B-figure. In the second session, an A-figure opposite in sex was paired subliminally with the same B-figure. Next, another B-figure was shown preceded by supraliminal exposure (.33 second) of another A-figure. Finally, ascending thresholds for the previously subliminal A-figures alone were also obtained, and *S*s gave associations to the symbolic A-figures. Control exposures without any preceding A-figures were also rated and drawn by all *S*s.

Effects of the subliminal figures were studied by comparing each *S*'s check-list rating of the B-figure with a group consensus of the appropriate A-figure. Drawings were rated for masculinity. Effects of the subliminal *genitals* were highly consistent on both check list and drawing, with *S*s tending either to incorporate or to exclude attributes of the subliminal A-figure in their impressions of the B-figure. Subliminal *symbols* were significantly effective in influencing *S*'s drawing of the B-figure; but this effect was not correlated with response on the check list. Recognition thresholds for the genital A-figures were related to check-list responsiveness, while thresholds for the symbolic A-figures were related only to responsiveness on the drawings.

[8] The fact that one *S*, whose data were subsequently discarded, did report awareness of "something else" preceding the picture, shows that *S*s could report such experiences if they had them.

REFERENCES

1. ADAMS, J. K. Laboratory studies of behavior without awareness. *Psychol. Bull.*, 1957, **54**, 383–405.
2. BINET, A. *Alterations of personality.* (Trans. by Helen G. Baldwin) New York: Appleton, 1896.
3. BLUM, G. S. Perceptual defense revisited. *J. abnorm. soc. Psychol.*, 1955, **51**, 24–29.
4. BRICKER, P. D., & CHAPANIS, A. Do incorrectly perceived tachistoscopic stimuli convey some information? *Psychol. Rev.*, 1953, **60**, 181–188.
5. CHEATHAM, P. G. Visual perceptual latency as a function of stimulus brightness and contour shape. *J. exp. Psychol.*, 1952, **43**, 369–380.

6. ERIKSEN, C. W. Perceptual defense as a function of unacceptable need. *J. abnorm. soc. Psychol.*, 1951, **46**, 557–565.
7. ERIKSEN, C. W. Subception: Fact or artifact? *Psychol. Rev.*, 1956, **63**, 74–80.
8. FISHER, C. Dreams and perception. *J. Amer. psychoanal. Ass.*, 1954, **2**, 389–445.
9. FISHER, C. Dreams, images and perception: A study of unconscious-preconscious relationships. *J. Amer. psychoanal. Ass.*, 1956, **4**, 5–47.
10. FREUD, S. *The interpretation of dreams.* Vols. IV & V. (Trans. by J. Strachey) London: Hogarth Press, 1953.
11. HELSON, H. Perception and personality—a critique of recent experimental literature. *USAF Sch. Aviat. Med. Proj. Rep.* 1953, Project #21-0202-0007 (Report #1).
12. KARLIN, L. The New York University tachistoscope. *Amer. J. Psychol.*, 1955, **68**, 462–466.
13. KRIS, E. On preconscious mental processes. *Psychoanal. Quart.*, 1950, **19**, 510–560.
14. LACEY, J. I., & SMITH, R. L. Conditioning and generalization of unconscious anxiety. *Science*, 1954, **120**, 1045–1052.
15. LAZARUS, R. S., & MCCLEARY, R. A. Autonomic discrimination without awareness: A study of subception. *Psychol. Rev.*, 1951, **48**, 113–122.
16. MCGINNIES, E. Emotionality and perceptual defense. *Psychol. Rev.*, 1949, **56**, 244–251.
17. MILLER, J. G. Unconscious processes and perception. In R. R. Blake & G. V. Ramsey (Eds.), *Perception: An approach to personality.* New York: Ronald, 1951.
18. OSGOOD, C. *The measurement of meaning.* Champaign: Univer. of Illinois Press, 1958.
19. POETZL, O. Experimentelle erregte Traumbilder in ihren Beziehungen zum indirekten Sehen. *Zfl. ges. Neurol. Psychiat.*, 1917, **37**, 278–349.
20. RAPAPORT, D. *Emotions and memory.* Baltimore: Williams & Wilkins, 1942.
21. SCHILDER, P. *Medical psychology.* (Trans. by D. Rapaport) New York: International Universities Press, 1953.
22. SHEVRIN, H., & LUBORSKY, L. The measurement of preconscious perception in dreams and images: An investigation of the Poetzl phenomenon. *J. abnorm. soc. Psychol.*, 1958, **56**, 285–294.
23. VOOR, J. H. Subliminal perception and subception. *J. Psychol.*, 1956, **41**, 437–458.
24. WALKER, H. M., & LEV, J. *Statistical inference.* New York: Henry Holt, 1953.
25. WERNER, H. Studies on contour: I. Qualitative analyses. *Amer. J. Psychol.*, 1935, **47**, 40–64.

Received July 26, 1957.

A STUDY OF ANXIETY AND LEARNING IN CHILDREN[1]

RICHARD R. WAITE, SEYMOUR B. SARASON, FREDERICK F. LIGHTHALL, AND KENNETH S. DAVIDSON

Yale University

THIS study focuses on differences between high anxious (HA) and low anxious (LA) children in an experimental learning situation. Two recent reports described test anxiety and general anxiety scales for children[2] and presented evidence relevant to the problem of validity (1, 6). The present study is similarly concerned with the validity of the children's scales. Do groups defined in terms of scores on the scales perform differently in an experimental learning situation? It was expected that any such over-all differences would be in the same direction as that found in college studies of test anxiety, LA subjects doing better than HA subjects (1, 2, 5). Another objective was the investigation of the effects of differential instructions (success, failure, and neutral) on HA and LA children. Again our expectations were derived from the results of the college studies in which LA subjects did best when given either success or failure instructions, whereas HA groups performed optimally under neutral conditions.

The conviction with which we hypothesized the above differences was tempered by a number of facts: (*a*) our present subjects were obviously much younger than college students, (*b*) our college students were males, whereas in the present study both sexes were used, and (*c*) the intelligence range of the college students was restricted, whereas in a grade school such limitations are much less. In the light of these and other considerations, it was felt that findings similar to those obtained with college students would not only be relevant to the validity of the scales but also to an understanding of the developmental aspects of the anxious reaction.

METHOD

Subjects

The children in this study came from a group of 747 pupils in Grades 2 through 5 to whom the test anxiety and general anxiety scales had been administered in the previous grade. All pupils had taken the Otis Alpha Intelligence Scale as well. Excluded were those regarded by their teachers as arithmetic, reading, or behavior problems, and those whose parents were separated or divorced. The *S*s whose test and general anxiety scores fell in the upper quartile of the distributions were considered HA, and those whose scores fell in the lower quartile were termed LA. Twenty-four pairs of children were then selected by matching LA *S*s with HA *S*s in terms of grade, sex, and Otis Alpha scores. Three such pairs were obtained for each grade-sex group, and one of these pairs was assigned to each of the three instruction groups—neutral, success, and failure. Thus, the design was a 2 (anxiety) by 3 (instructions) by 2 (sex) by 4 (grade) arrangement, with one *S* in a cell. The matching procedure, even with such a large group of *S*s from which to draw pairs, was found to be difficult. It was evident that to have had more than one person in a cell, while at the same time including all the specified variables, would have required many more than the original 747 children.

Procedure

The *E*, who did not know which *S*s were classified as HA or LA, worked with them individually in the public schools. After bringing *S* from the classroom to the testing room, an introductory conversation took place during which the establishment of good rapport was attempted. Then *E* asked the *S* to tell a story to a TAT card,[3] after which the following instructions were given.

> I have a test I'd like you to take today. This is a test that I have given to lots of children in all of the schools in Milford. I've given the test to children in all the grades, so that I know just how well a first grader can do, how well a second grader can do, a third grader, a fourth, fifth, sixth, and seventh grader can do. Of course, a second grader does better than a first grader, a fourth grader does better than a third grader, and so on, because he's older, right? Now, your teacher and I want to find out how well *you* can do on the test—whether you do as well as other [4th] graders, or only as well as [3rd or 2nd] graders, or as good as [5th or 6th] graders.
>
> Now, I have 12 cards here with a figure and a number on each. There's a different figure and a different number on each card. I'm going to show you the cards one at a time, and I want you to try to remember the number that goes with each figure. When I've shown you all the cards once, I'll give you a piece of paper with the figures on it and you try to write the correct number under each of the figures. Then I'll show you the cards again and ask you to

[1] The project of which this study is a part is being supported by a grant from the United States Public Health Service, (M-712).

[2] Copies of these two scales are available for research purposes only.

[3] We wish to express our thanks to Robert P. Abelson for his most helpful advice on the statistical analyses.

write the numbers again. We'll do this six times in all, so you'll have plenty of chances. O.K.?

The cards were presented at two-second intervals, and *S*s were allowed as much time as they desired in writing the numbers. While they were filling out the "answer sheets" for each trial, *E* pretended that he was correcting the results of the previous trial. Upon completion of the sixth trial, *E* "added up the results" and then made the following statements, appropriate to the instruction group of which the *S* was a member.

Success Ss. "You did as well as kids in the [2 grades above *S*'s actual grade] grade. You can be very proud. Now I'm going to give you another test just like this one. This one has the same numbers, but different figures."

Failure Ss. "You didn't do well at all on this test. You only did as well as [2 grades below *S*'s actual grade] graders. You should have done much better. Now I'm going to give you another test just like this one. This one has the same numbers, but different figures."

Neutral Ss. "Now I'm going to give you another test just like this one. This one has the same numbers, but different figures."

The second set of six trials was then given, using a new set of stimulus cards. Then all *S*s were assured that they had done well, with extra attention being given those children who had received failure instructions, in order to leave them at ease. After *S* had returned to his classroom, *E* filled out a behavioral check list which included items such as nail-biting, scratching, apparent motivation, reaction to instructions, etc., and wrote a brief report on his general observations during the testing situation.

RESULTS

Figure 1 presents the over-all results of the two anxiety groups. For each group, mean scores on each trial were calculated and plotted to provide group learning curves. At the beginning of the first task, the group means

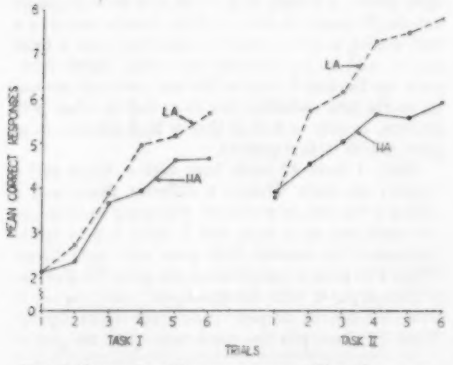

FIG. 1. LEARNING CURVES FOR THE TWO ANXIETY GROUPS BEFORE AND AFTER INSTRUCTIONS

are identical, but by the conclusion of this task, the LA group is decidedly superior in performance. When the same data are divided into sex-anxiety groups, the differences in anxiety-group means hold up better within males than within females. In general, the girls appear to do better than the boys.

Individual scores were treated in the following manner: First, for each *S* a mean score over each set of six trials was computed. Second, the score of the first trial in each of the two sets of trials was multiplied by -5, the second by -3, the third by -1, the fourth by 1, the fifth by 3, and the sixth by 5. These products were then summed to give a comparison number which can be considered an index of the linear component of the slope of the individual's learning curve, or an estimate of his rate of learning. For each *S*, then, there are two mean scores (one for each set of six trials) and two slope scores.

Since there were no differential instructions given on Task I, it was possible to carry out an analysis of variance on the first set of trials, collapsing on instructions to leave three *S*s in a cell. Two analyses were done, one for slopes and one for means. In terms of means, there is no significant difference between the LA and HA groups. The grade effect is significant, but the data show that instead of an increase in mean scores from grade to grade as we would expect, there is a drop from second to third grade, and then an increase to fourth and to fifth grade. The analysis of slopes shows parallel results. See Footnote 3.

Next, analyses of the Task II data were carried out. In this case it was necessary to construct an error term from the grade-by-

TABLE 1
ANALYSIS OF VARIANCE OF TASK II MEAN SCORES

Source	*df*	Mean Square	*F*	*p*
Sex	1	18.54	4.14	.10
Instructions	2	.40	—	—
Grade	3	50.85	11.43	.01
Anxiety	1	18.13	4.08	.10
S × I	2	4.06	—	—
S × G	3	10.42	2.34	.10
S × A	1	1.63	—	—
I × A	2	4.25	—	—
G × A	3	14.76	3.32	.05
S × I × A	2	1.15	—	—
S × G × A	3	3.55	—	—
Error	24	4.45		

instructions interactions; grade was used because it was thought to be a relatively unimportant variable in terms of the purposes of this study. The analysis of means (Table 1) again shows significant grade effects. Anxiety and sex are significant at the .10 level, while the grade-by-anxiety interaction is significant at the .05 level. The data show that the girls did somewhat better than the boys, that the grade effect is essentially linear, and that LA *Ss* did better than HA *Ss*. Furthermore, the significance of the grade-by-anxiety interaction seems due, to a great extent, to the first-grade HA girls scoring much higher than one would expect (higher than fourth-grade HA or LA girls). The analysis of slopes (Table 2), on the other hand, gives much clearer results. Only anxiety is significant and here at the .01 level. LA *Ss* learned faster than HA *Ss*. Of note is the nonsignificance of the instructions effect in both analyses.

Since the differences in slopes of the anxiety groups may have been a carry-over from the first task, sufficiently increased through practice to reach significance, an analysis of covariance was performed on the Task II slopes, controlling for Task I slopes. Table 3 shows that the difference in Task II slopes remains significant at the .05 level.

The check lists which were obtained at the end of each experimental session were totalled for each anxiety group. No differences in any of the items were found.

TABLE 2

ANALYSIS OF VARIANCE OF TASK II SLOPE SCORES

Source	*df*	Mean Square	*F*	*p*
Sex	1	295.02	—	—
Instructions	2	207.06	—	—
Grade	3	569.35	1.54	—
Anxiety	1	2041.02	5.52	.01
S × I	2	116.90	—	—
S × G	3	246.58	—	—
S × A	1	792.19	2.14	—
I × A	2	585.27	—	—
G × A	3	299.58	—	—
S × I × A	2	324.19	—	—
S × G × A	3	341.19	—	—
Error	24	369.22		

TABLE 3

ANALYSIS OF COVARIANCE OF TASK II SLOPE SCORES

Source	*df*	Adjusted Mean Square	*F*	*p*
Anxiety	1	1255.98	11.33	.05
Within	45	110.87		
Total	46			

DISCUSSION

The most interesting results as far as the major purpose of the experiment is concerned are the significant differences between the anxiety groups. Since the groups were at least roughly matched on other factors which operate in learning situations with children—sex, age, and intelligence—we can only assume that whatever it is that the anxiety scales measure is highly related to the differences in performance. As expected, the direction of the results coincides with that obtained in similar studies with college students. Thus, we can conclude that the experiment has further demonstrated the validity of the children's scales.

The lack of any significant difference between the anxiety groups on the first task is an aspect of the results about which we can only conjecture. One explanation may lie in the nature of the two learning tasks. While they were alike in manner of presentation, size, other formal characteristics of the stimuli, and the nature of the learning involved, there was a slightly higher intralist similarity on the first task. Consequently, it would be expected that the learning involved on the first task would be more difficult, both because of the higher intralist similarity and the initial position of the task. The net result may be thought of in terms of a general depression of scores, cutting down the possibility of differences.

An alternative explanation involves viewing the effect of instructions in a different way. In this study it apparently made no difference what was said to the *Ss*. Perhaps it was the case that any statement about another test, no matter what the specific verbalization, cued off different reactions in the two anxiety groups, reactions which ultimately affected performance. The present study provides no information relevant to the veracity of this hypothesis, however.

One other problem deserves mention and brief discussion. One of our interests in anxiety, particularly test anxiety, concerns both its interfering and facilitating characteristics in

different kinds of learning situations. In the present study, the differences between HA and LA *S*s suggest the presence of interfering factors in the former group. Whatever the degree of interference, its significance derives in large part from the fact that the two groups were matched in terms of an intelligence test score, the assumption underlying such matching being that the two groups are approximately equal in general intelligence and should therefore perform similarly in the learning situation. However, since our studies of children and college students (4) have consistently a significant *negative* relationship between IQ score and test anxiety,[4] it should follow that matching HA and LA *S*s on IQ, as was done here, involves matching on a measure which is already influenced by anxiety. It could be argued, therefore, that when HA and LA groups are so matched, the "true" score of the former is likely to be higher than that of the latter. From this standpoint, one may speculate that the performance of the HA *S*s in the present study reflects a degree of interference that is greater than a comparison of the learning curves of HA and LA *S*s indicates; i.e., the degree of discrepancy between potential and performance is greater among HA *S*s than the learning curves suggest.

[4] We have computed correlations between IQ and test anxiety for several hundred classrooms. Approximately 90% of these correlations were negative.

SUMMARY

A general anxiety scale and a test anxiety scale were administered to 747 grade school children. Out of this group, 24 HA–LA pairs of subjects were given two modified paired-associate learning tasks, separated by neutral, failure, and success instructions. The results showed no differences due to instructions, but LA subjects performed better than HA subjects on the second task. Alternative explanations for the lack of differences on the first task were offered. The study was interpreted as positive evidence for the validity of the anxiety scales.

REFERENCES

1. GORDON, E. A. Self esteem, need for achievement, test anxiety and performance. Unpublished doctoral dissertation, Yale Univer., 1955.
2. MANDLER, G., & SARASON, S. B. A study of anxiety and learning *J. abnorm. soc. Psychol.*, 1952, **47**, 166–173.
3. SARASON, S. B., DAVIDSON, K. S., LIGHTHALL, F. F., & WAITE R. R. A test anxiety scale for children. *Child Develpm.*, in press.
4. SARASON, S. B., & MANDLER, G. Some correlates of test anxiety. *J. abnorm. soc. Psychol.*, 1952, **47**, 810–817.
5. SARASON, S. B., MANDLER, G., & CRAIGHILL, P. G. The effect of differential instructions on anxiety and learning. *J. abnorm. soc. Psychol.*, 1952, **47**, 561–565.
6. SARNOFF, I., LIGHTHALL, F. F., WAITE, R. R., DAVIDSON, K. S., & SARASON, S. B. A cross-cultural study of anxiety amongst American and English school children. *J. educ. Psychol.*, 1958, **49**, 129–136.

Received July 29, 1957.

THE EFFECTS OF DIFFERENT SOCIAL FEEDBACK CONDITIONS UPON PERFORMANCE IN DYADIC TEAMS[1]

SEYMOUR ROSENBERG[2] AND ROBERT L. HALL[2]

USAF Personnel and Training Research Center

IN THE present study, one of three types of feedback stimuli was presented to a group member following a response. They are: (*a*) information based upon the error magnitude of the subject's own response only, termed "direct" values. This type of feedback stimulus is typically used in individual learning studies to differentiate a particular range of response magnitudes; (*b*) information that combines in a simple additive function the error magnitudes of both the subject's own response and that of his teammate, termed "confounded" values; (*c*) information based upon the error magnitude of a teammate's response only, termed "other's" values. These three feedback conditions may be viewed as comprising two extremes and an intermediate point of a continuum, which specifies the fractional weight given to a subject's response relative to the fractional weight given to his teammate's response in the feedback stimulus to the *subject*. The feedback value to the subject is "direct" when his response weight in the feedback is 1.0 and his teammate's response weight is 0.0 (one extreme of the continuum). The subject's feedback value is "other's" when his own response weight in his feedback is 0.0 and his teammate's is 1.00 (other extreme of the continuum). The subject's feedback value is "confounded" when his feedback values contain nonzero weighing of both his own and his teammate's response. The present paper will describe the effects of these feedback conditions upon individual and team measures, and upon changes in these measures over a series of discrete trials.

The stimulus event following a response has been variously termed: feedback, knowledge of performance, knowledge of results, reinforcement, terminating stimulus, etc. (e.g., 2, 3, 9, 11). There is little disagreement about the importance of such general parameters of this stimulus event as its frequency and periodicity, coarseness of measurement, delay, manner of presentation, type and amount of informational error, number of trials, etc. However, these parameters have been studied almost exclusively in experiments using individual subjects. Since there has been very little systematic attempt either to determine the values of such parameters particularly relevant to interpersonal settings, or to formulate variables unique to these settings, the research to date is incomplete or, at best, difficult to apply to interpersonal settings. The application or extension of parameters to interpersonal contexts presents a number of interesting and novel problems. For example, when a stimulus event following a response by Person A is some joint function of his behavior and that of a Person B, it is difficult to specify in quantified terms the stimulation of A from knowledge obtained in individual research and, further, to specify changes in stimulus parameters that occur over time. In attempting to apply current knowledge to interpersonal settings, the investigator may also encounter the need to formulate novel variables. For example, one class of variables unique to interpersonal settings is variations in the mathematical function that expresses the way the responses of Persons A and B are combined to produce a stimulus event to one or both persons. Distributional variables of any sort, i.e., combinations of treatments of group members, are also novel. Neither in social psychology nor in learning psychology has there been any systematic experimental analysis of social and learning variables simultaneously. This study is one of the early experiments in such a program. This program was described in some detail in a paper by Hall and Rosenberg.[3]

[1] This report is based on work done under ARDC Project No. 7713, Task No. 27104, in support of the research and development program of the Air Force Personnel and Training Research Center, Lackland Air Force Base, Texas. Permission is granted for reproduction, translation, publication, use, and disposal in whole and in part by or for the United States Government.

[2] Seymour Rosenberg is now at the University of Kansas. Robert L. Hall is now at the University of Minnesota.

[3] Hall, R. L., & Rosenberg, S. An S-R analysis of social interaction. Paper presented at American Psychological Association meetings in New York, September 1957.

METHOD

Subjects. The subjects (*Ss*) were 54 male Air Force basic trainees whose modal age was about 17 years. They were assigned as a unit for use as *Ss* in psychological research. *Ss* from such units were excluded from the experiment if their Armed Forces Qualification Test (AFQT) score was below the 65th percentile rank, which is roughly equivalent to a Stanford Binet IQ of 108 (4). *Ss* were used in pairs, randomly assembled.

Apparatus. The two *Ss* were seated in identical adjacent booths. The walls of the booth were sufficiently high so that *Ss* could not see each other. One wall of each booth consisted of the panel and table shown in Figure 1. Each *S* faced a separate panel. A micrometer was mounted behind the shield so that the knob at the beginning of each trial projected ¼ inch from the face of the panel. The *S* could touch only the knob portion of the micrometer. He could turn the knob by reaching under the shield, but he received no visual cues nor special tactual cues from the knob to indicate the amount it was turned. The knob could only be turned to the left. The experimenter (*E*), seated on the other side of the panel, could read the micrometer scale, which is a linear measure of the amount the knob is turned. Also, on the panel in front of each *S*, there was a linear scale extending from zero in the center to 80 in each direction, indicating "size of error." A movable pointer above the scale displayed a score to the *S* after each trial. The scale was concealed by a hinged shield during the *Ss*' response period and the feedback delay period. The scale also contained "off the scale" marks at each end. At the center of the scale an interval of 10 units was marked off by two red lines and was labeled "try to get pointer within red lines for good score." No cues were present on the scale to indicate whether error scores to the left of zero on the scale indicated too large or too small a response magnitude. For counterbalancing, excessive response magnitudes were shown to the left of zero for half of the *Ss* and to the right of zero for the other half of the *Ss*.

Procedure. Each pair of *Ss* was given the same instructions, followed by 50 response-feedback trials. Each member of a pair was independently assigned to one of the three feedback conditions for all 50 trials. To satisfy a convenient analysis of variance design to be described later, certain restrictions were imposed upon the numbers of *Ss* assigned to each condition.

After the two *Ss* were seated in adjacent booths, a few minutes were spent by the *E* "setting them at ease." The identical panels, one for each *S*, were then put in place. The following tape-recorded instructions were played:

> The Air Force needs to know more about how well people can do certain tasks. The experiment in which you are to take part is a study of just one such task. It is a simple task of turning a knob the right amount. Listen carefully to these instructions and the experimenter will point out things as they are mentioned. For this experiment a panel will be placed beside each of you. On the front of the panel you see a shield and underneath the shield, where you cannot see it, is a small knob. Do not look under the shield. You are expected to work without seeing the knob. Your job is to learn to turn your knob just the right amount. You will have a number of chances to do this. To help you, you will be shown, after each trial, the error in the score that you made. Your error score will be shown you on the scoreboard at the top of your panel. The movable pointer, which is moving now, tells you how much you missed. To make a *good* score, try each time to get the pointer within the two red lines. Always turn your knob to the left as shown by the small sign tacked on your panel. A buzzer will be used to tell you when to start and stop. The buzzer sounds like this. [Buzzer is sounded.] When the buzzer starts, you start to make your setting. As soon as you have made the setting you think is correct, take your hand away from the knob. When the buzzer stops, you must take your hand away whether you are finished making your setting or not. You will find that you have plenty of time without rushing. A few seconds after the buzzer stops you will be shown on your scoreboard the error in the setting you just made. Shortly after that, you will hear the buzzer again. This means that you are to start again and try once more to make a good score. After each trial your knob is set back by us. Each time the buzzer starts, you start over and try again to make a good score. Now we are ready to start. Do not touch your knob until you hear the buzzer.

It should be noted that the *Ss* were given no cues that they might be working as a team. They did not talk to each other. To control for inadvertent cues by *E* to *Ss* in different treatments, each pair of *Ss* was assigned to one of the treatments only *after Ss* were seated in the closed booths and instructed.

The procedure for one complete trial consisted of three automatically timed phases: (*a*) a response period of eight seconds, indicated to the *Ss* by a mild buzzer, during which the knob could be turned; (*b*) a delay of eight seconds during which *E* recorded response magnitude and moved the scale pointer (still concealed from each *S*) to its appropriate value for that trial; (*c*) exposure of the pointer position to the *S* by lifting the hinged shield. During this interval the micrometer was reset by *E* to its original position. The shield was

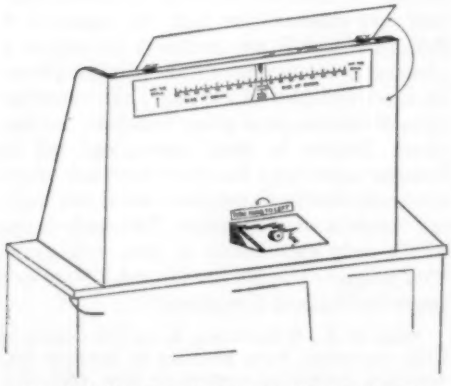

FIG. 1. A SKETCH OF THE PANEL FACING EACH *S*

TABLE 1

SUMMARY OF THE STATISTICAL ANALYSES OF THE THREE RESPONSE MEASURES

Source	*df* of each Analysis			Individual Accuracy ($A_i = \log(\lvert X - 100 \rvert + 1)$)		Team Accuracy ($A_t = \log\lvert X + Y - 200 \rvert + 1$)		Differentiation ($D_t = \log(X - Y) + 1$		Error Term for *F*
				Mean Sq.	*F*	Mean Sq.	*F*	Mean Sq.	*F*	
1. Teams	26			83009		70602		59782		
A. Rows (Feedback cond.)		2		620276	17.32***	174340 [a]	5.46**	209026 [a]	3.46*	1D
B. Columns (Feedback cond.)		2		10784	.30	284080 [a]		121978 [a]		1D
C. Rows × Columns		4		63434	1.77	40915	.98	7984	.17	1D
D. Within Cells		18		35817		41952		47799		
2. Trials	49			2312		4102		3041		
3. Teams × Trials	1274			1405		1559		1559		
A. Rows × Trials		98		2089	1.54**	1242	1.09 [a]	1439	.89 [a]	3D
a. Linear			2	22418	4.51*	3854	2.38	16513	.92	3Da
b. Quadratic			2	10372	4.19*	4438	1.56	2452	1.09	3Db
c. Residual			94	1480	1.18	1119	.90	1097	.87	3Dc
B. Columns × Trials		98		1025	.75	2120	— [a]	1515	— [a]	3D
a. Linear			2	1274	.26	35148	— [a]	5046	— [a]	3Da
b. Quadratic			2	500	.20	5582	— [a]	6890	— [a]	3Db
c. Residual			94	1031	.82	1343	— [a]	1325	— [a]	3Dc
C. Rows × Columns × Trials		196		1463	1.08	1486	.96	1448	.87	3D
a. Linear			2	6751	1.36	7806	.95	9863	.84	3Da
b. Quadratic			2	10263	4.15*	2706	.84	6140	1.43	3Db
c. Residual			94	1163	.92	1326	.97	1169	.84	3Dc
D. Within Cells × Trials		882		1358		1548		1660		
a. Linear			18	4966		8201		11753		
b. Quadratic			18	2473		3212		4280		
c. Residual			846	1258		1371		1389		

[a] For each of the two team measures, A_t and D_t, rows and columns mean squares refer to exactly the same effects. Therefore, these two sources of variance were pooled within each analysis of variance and a single *F* ratio computed. The same procedure was adapted for the "rows × trials" and "column × trials" variances. Note that the numerator of such *F*s has associated with it four *df*.

*$P < .05$; **$P < .01$; ***$P < .001$.

lowered after 5–10 seconds and the response interval for a new trial was started.

Response magnitude for each *S* during each trial was read by the *E* directly from the micrometer scale. Twenty-five scale units equals one complete knob rotation. Where an *S* was receiving "direct" feedback (his response magnitude only), a perfect score (goal score) was 100 micrometer units (four knob turns). Where an *S* was receiving "confounded" feedback (the combination of his own and his teammate's response magnitude), a score to the *S* was computed by averaging the response magnitudes of the two persons. An average of 100 units was a perfect score. An *S* receiving "other's" feedback (his teammate's response magnitude) received a perfect score when his teammate made four knob turns or 100 micrometer units.

Measurement and statistical design. Three measures were computed for each trial and transformed to logarithms to reduce skewness:[4] (*a*) "individual accuracy," $A_i = \log(\lvert X - 100 \rvert + 1)$, where X is response magnitude, in micrometer units of one *S*; (*b*) "team accuracy," $A_t = \log(\lvert X + Y - 200 \rvert + 1)$, where X and Y are response magnitudes of the two *S*s in micrometer units; (*c*) "response (role) differentiation," $D_t = \log(\lvert X - Y \rvert + 1)$.

Each of the three measures was subjected to a similar analysis of variance. The design consists of three rows corresponding to the three feedback conditions to which one *S* could be assigned and three columns corresponding to the same three feedback conditions to which his teammate could be assigned. Thus, there are nine row-column cells. Three pairs of *S*s were assigned to each of these nine cells (combinations of feedback conditions). Fifty trials constitutes a third criterion of classification in the analysis.

For the A_i measure, one member of each dyad was randomly selected. Data from the second member were not analyzed. Thus, the same team is never represented twice in the analysis of A_i. Two analyses of variance of A_i were also avoided because of the complex experimental dependency between the two *S*s in most of the combinations of feedback conditions. The rows represent the feedback condition of the member whose A_i is contained in the cell. The columns represent the condition of his partner.

When either team measure (A_t or D_t) is used in this statistical design, row and column classifications bear the same relationship to the dependent variable, i.e., represent the feedback condition for one of the team members. Put in another way, the row variable is exactly the same as the column variable. In such a design, combinations of experimental conditions on one side of one of the diagonals are also found on the other side of that diagonal. Data must be arbitrarily distributed between cell-pairs whose conditions are the same.

Two orthogonal polynomial components were extracted from interaction effects containing trials. The

[4] The value, 1, was added to the absolute value in each of the three measures to avoid the log of 0, which has an indeterminate value.

components were the "linear" and "quadratic" and were used to test the presence of differential rates of development of a measure with trials. These components supplement the test for over-all trials effects since the latter is too general a test of trend and fails to distinguish simple trend types. The rationale and method of extraction of a linear component in a trend analysis is described in several places in the psychological literature (1, 8), and the more generalized procedure for extraction of orthogonal polynomials is described by Grant (5).

RESULTS

Analyses of variance of the three measures are summarized in Table 1. Note that row and column effects are pooled in the analysis of A_t and D_t since the effects tested are the same. For A_t, the effects refer to two different operations with respect to the dependent variable, i.e., an *S*'s response measure classified by his own feedback conditions (rows) and that of his partner (columns).

In the case of A_i, different feedback conditions of the *S*s result in significant differences in average level of his accuracy over the 50 trials (row effects), in over-all trend differences (interaction of rows with trials), and in linear and quadratic trend components. Figure 2(A) illustrates these differences. Note that on the ordinate, the smaller A_i values are associated with greater accuracy, and that on the abscissa the trials are grouped in sets of five. The graph points up a rather simple positive relationship between individual accuracy and contribution of the *S* to his own feedback. In striking contrast, different feedback conditions of *S*'s partner (column effects) do not result in any significant differences in *S*s' accuracy. This is apparent from Figure 2(B) which illustrates no consistent divergencies among groups. Figures 2(A) and 2(B) would not illustrate the presence of any unique effects produced by particular combinations of an *S*'s condition and that of his partner, if any existed. For example, if the combination of "other's-other's" conditions resulted in an exaggerated degree of inaccuracy for *S* (over and above that accounted for by the main effects), these graphs would fail to show this effect. However, such an effect would be reflected in "rows × columns" and/or "rows × columns × trials" interactions in the A_i analysis. The results in Table 1 fail to support, with any statistical significance, the presence of such combinatorial effects except in the isolated instances of a quadratic component in Source 3C.

An examination of both team measures, A_t and D_t, reveals a rather consistent picture. For both measures, differences in the feedback conditions yield significant differences on the average of the 50 trials (Sources: Sum of 1A and 1B). Differences in trend, however, are not revealed (Sources: Sum of 3A and 3B) unless one is willing to use a more liberal confidence level (or another error term) as a basis for rejecting null hypothesis. In the case of A_t, the *F* value of 2.38 (*df*, 4, 18) for linear differences in trend falls between the .05 and .10 levels. Figures 3(A) and 3(B) illustrate the A_t data for rows and columns, respectively, and do suggest the existence of trend differences among conditions. The "replication" between row effects and column effects does not appear to be perfect, but no reversals are apparent. It is interesting to note, in Figure 3(B), that confounded feedback results in somewhat bet-

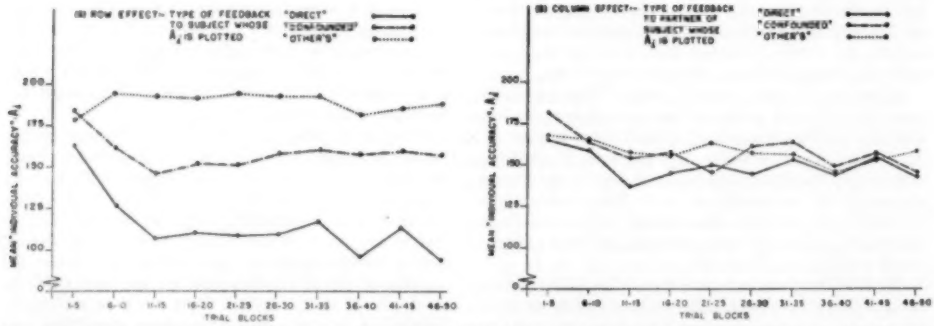

FIG. 2. DIFFERENCES AMONG GROUPS OF *S*s IN THEIR MEAN A_i AS A RESULT OF DIFFERENCES IN: (*a*) *THEIR OWN* FEEDBACK CONDITIONS; (*b*) FEEDBACK CONDITIONS OF THEIR *PARTNERS*

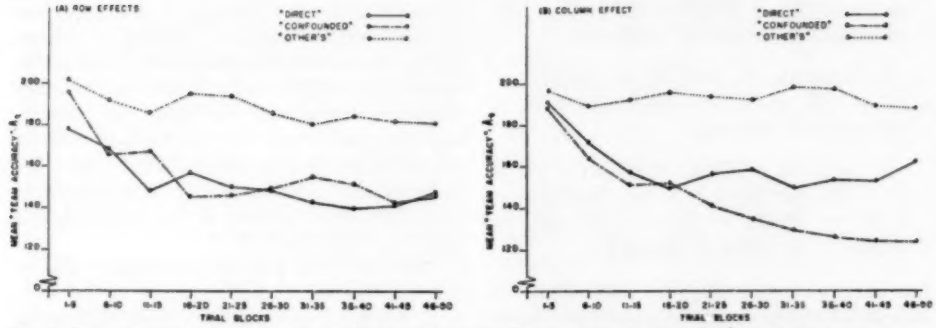

FIG. 3. DIFFERENCES AMONG GROUPS OF SUBJECT-PAIRS IN THEIR MEAN A_t AS A RESULT OF DIFFERENCES IN FEEDBACK CONDITIONS TO: (*a*) ONE MEMBER OF THE PAIR; (*b*) THE SECOND MEMBER OF THE PAIR

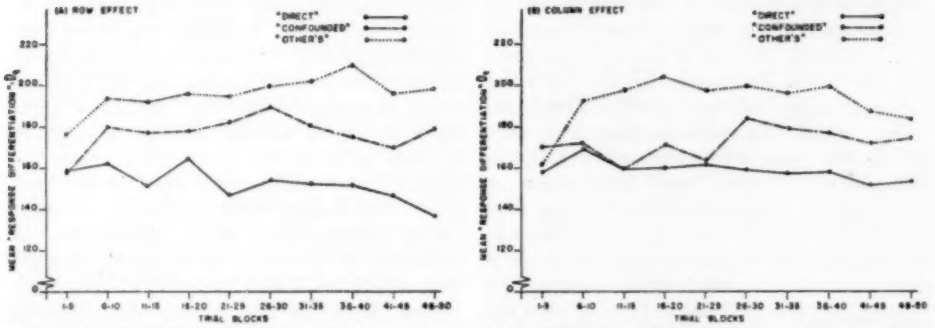

FIG. 4. DIFFERENCES AMONG GROUPS OF SUBJECT-PAIRS IN THEIR MEAN D_t AS A RESULT OF DIFFERENCES IN FEEDBACK CONDITIONS TO: (*a*) ONE MEMBER OF THE PAIR; (*b*) THE SECOND MEMBER OF THE PAIR

ter team accuracy than direct feedback to each *S*. Figures 4(A) and 4(B) present for rows and columns, respectively, the D_t measure, the ordinate indicating the mean magnitude of difference between teammates' responses. "Replication" is again somewhat imperfect although no serious contradictions are apparent. Teams containing an *S* who received as feedback only his teammate's response showed greatest D_t, while teams with an *S* receiving "direct" feedback showed the least D_t.

Finally, it is again important to note that Figures 3 and 4 would not illustrate unique effects in A_t and D_t produced by particular combinations of feedback conditions to which pairs of *S*s were exposed. However, the analyses of the two team measures in Table 1 fail to support the statistical significance of any such combinational effects since none of the "rows × columns" or "rows × columns × trials" interactions are significant. Team accuracy and response differentiation values (as quantified in this study) are apparently determinable from an additive weighting of each of the two feedback conditions of the *S*s.

DISCUSSION

As noted previously, the three feedback conditions in the study can be viewed as points on a continuum. The continuum is describable as the proportion of contribution of an *S*'s own response to his feedback, relative to the contribution of his teammate's response. The continuum may be expressed as the ratio of the weights accorded responses of the two *S*s in a simple weighted average; that is, the size of the ratio, k_1/k_2, in the simple function:

$$\text{Feedback} = k_1X + k_2Y - G$$

where X and Y are response magnitudes of two *S*s, G is goal score, and $k_1 + k_2 = 1$. The three feedback conditions would then be expressed:

a. Feedback = (1)(X) + (0)(Y) − 100 = X − 100, for Subject X, or Feed-

back = (0)(X) + (1)(Y) − 100 = Y − 100, for Subject Y, in "direct" feedback.

b. Feedback = (.5)(X) + (.5)(Y) − 100 = .5X + .5Y − 100 for both *S*s, in "confounded" feedback.

c. Feedback = (0)(X) + (1)(Y) − 100 = Y − 100, for Subject X, or Feedback = (1)(X) + (0)(Y) − 100 = X − 100, for Subject Y, in "other's" feedback.

The findings concerning A_i (individual accuracy) are rather simply summarized in terms of its relation to the k_1/k_2 ratio. Both average magnitude and rate of development of A_i are related monotonically to the magnitude of k_1/k_2 where k_1 and k_2 are the constants applied to the feedback function of the *S*. No relationship was found between A_i of an *S* and the k_1/k_2 ratio of his teammate's feedback function.

The use of this ratio to summarize succinctly the relationships discovered early in this research program imply little about its future utility. A more articulated theory may not use this variable. One could also describe, for example, the feedback conditions along a continuum or continua which express the "quality of the feedback." Quality of feedback may refer to the size *and* consistency of the covariance between feedback values and the response magnitudes of an *S* (i.e., reliability of feedback as a measure of response magnitude). The correlation coefficient may be used as a measure of this covariance, although more comprehensive measurement techniques are probably required.[5] Quality of feedback may also refer to any constant difference in magnitude between a feedback value and a response value. In other words, quality is reduced by both variable and constant errors in the feedback. It may be noted that "quality of feedback" is oriented toward individual feedback. Such an orientation points up, however, a number of interesting feedback variables that are likely to arise in interpersonal settings. For example, there is some reason to expect that the quality of feedback to an *S* may change systematically with time. In confounded feedback, the variable error in the feedback with respect to the response probably decreases with time (as each *S* "learns"), but constant errors stabilize. Individual learning experiments have generally tended to maintain such conditions constant over trials rather than to systematically vary them. An interesting exception occurs where a response is being gradually molded or differentiated. Here the R-S relationship is manipulated as a function of trials (or a joint function of trials and response adequacy) (11, chap. 8).

The two team measures, especially A_t, do not bear the same relationships with the feedback condition as does A_i. Indeed, differential trend effects are difficult to establish at all except through the use of liberal confidence levels. The difference in average value of A_t among the three feedback conditions are highly significant, however, and are worthy of comment because of an apparent difference in result as compared with A_i. Confounded feedback to a team member appears to result in at least as accurate, and possibly more accurate, team performance than direct feedback. This finding, if valid, points up an interesting phenomenon about distribution of feedback to teams where a group product rather than an individual product is the criterion. Since the task is one which permits compensation by one team member for difficult-to-extinguish constant errors of the second, dyads, *on the average*, would yield superior group products. The D_t findings attest to a greater difference in response magnitude between team members in confounded feedback than in direct feedback. This difference persists over trials and is consistent with the above hypothesis but does not obviate other explanations. Indeed, another tenable hypothesis is that the feedback values displayed to an *S* are much finer (more detailed) than he can discriminate kinesthetically. The inappropriate detail then acts as "noise" in the feedback producing response variability. In the confounded feedback condition, the feedback values (although as detailed as in direct feedback) are less variable from trial to trial than in direct feedback, since the former are based on an average of two response magnitudes. *S*s, in turn, may respond with less variability *about the goal score* from trial to trial. Further study of the phenomenon would be necessary both to replicate the differences in the A_t found in the present study and to examine intentionally the hypotheses stated here.

[5] Some preliminary attempts in this direction were made with these data, but the problems remain sufficiently horrendous to prevent adequate report of these at the present time.

The D_t measure, aside from corroborating the hypothesis above, has special interest for social psychologists. The response distribution in an interpersonal relationship receives considerable attention under such headings as role, role differentiation, role distribution, etc. (e.g., 6, 7, 10). The D_t measure is a considerable simplification of the usual role differentiation measures, but seems to represent the same type of measurement operation. The simplicity stems from the fact that a single response class (i.e., a single dimension of behavior) is used in this study. The analysis of the D_t measure under different feedback conditions points up a rather novel independent variable for social psychology. Response or role differentiation, D_t, bears the same simple relationship to the k_1/k_2 ratio as does A_t. The trend relationships are not as apparent, however.

The D_t measure also has some interesting implications for team reassemblies or team-member interchangeability. Conditions, whether they be feedback or other, which minimize the D_t value would appear to provide greatest flexibility for team reassemblies since teammates who respond similarly can be reassembled with a minimum of training. Teams requiring differentiation among a number of response classes for effective performance may require other feedback conditions.

SUMMARY

Knowledge of results (feedback) available to an individual working in a team may be a measure of: (*a*) his own response only, (*b*) a combination of his own and a teammate's response, (*c*) a teammate's response only. Each teammate may receive any of these three types of feedback.

The independent variable consisted of these three conditions of feedback in all six combinations. *S*s were used in pairs, and each pair was given 50 trials on a simple knob-turning task with one of these combinations of feedback conditions. A total of 27 pairs of *S*s was used in the study.

Three response measures were analyzed: (*a*) individual response accuracy, (*b*) team accuracy, (*c*) response (role) differentiation between team members. Analyses revealed significant differences between feedback conditions on each of these measures. Teams containing an *S* who received as feedback only his teammate's response showed poorest individual and team accuracy and greatest response differentiation. Differences among conditions in trend were consistently significant only for the individual accuracy measure. Learning rate on this measure was best when an *S* received information about his own response only and poorest when information consisted only of his teammate's response.

REFERENCES

1. ALEXANDER, H. W. A general test for trend. *Psychol. Bull.*, 1946, **43**, 533–537.
2. ANNETT, J., & KAY, H. Skilled performance. *Occup. Psychol.*, 1956, **30**, 112–117.
3. BROWN, J. S. A proposed program of research on psychological feedback (knowledge of results) in performance of psychomotor tasks. In R. M. Gagne (Ed.), *Research planning conference on perceptual and motor skills. USAF, ATRC, HRRC Conference Rep.* 49-2. San Antonio, Tex. Lackland AFB, 1949.
4. GRAGG, D. B., & DOUGLASS, H. J. Conversion tables for selected airman classification battery scores and comparable scores on other selected service and civilian tests. *USAF Pers. Train. Res. Cent. Res. Rep.*, 1956, No. AFPTRC-TN-56-89.
5. GRANT, D. A. Analysis-of-variance tests in the analysis and comparison of curves. *Psychol. Bull.*, 1956, **53**, 141–154.
6. HARE, A. P., BORGATTA, E. F., & BALES, R. F. (Eds.) *Small groups*. New York: Alfred A. Knopf, 1955. Chap. 9.
7. HOMANS, G. C. *The human group*. New York: Harcourt, Brace, 1950.
8. LINDQUIST, E. F. Goodness of fit on trend curves and significance of trend differences. *Psychometrika*, 1947, **12**, 65–78.
9. MILLER, R. B. *Handbook on training and training equipment design. USAF WADC Tech. Rep.*, 1953, No. 53-136.
10. NEWCOMB, T. M., & HARTLEY, E. L. (Eds.) *Readings in social psychology*. New York: Henry Holt, 1947. Part VIII.
11. SKINNER, B. F. *The behavior of organisms*. New York: Appleton-Century, 1938.

Received July 29, 1957.

A COMPARISON OF INTERVIEW AND OBSERVATION MEASURES OF MOTHER BEHAVIOR[1]

HENRIETTA T. SMITH

Vassar College

THE study of parent-child relationships has been largely dependent upon indirect measures of behavior, including such techniques as rating scales (1), questionnaires, interviews, and case histories. Merrill (4) found that controlled observations of interactions between mothers and children could be used to isolate and correlate significant aspects of behavior with a high degree of reliability. The present study is an attempt to evaluate the adequacy of interviews as compared with such controlled observations of actual behavior.

The interview method is economical for it enables the investigator to collect data on past and present maternal attitudes and behavior, and it allows comparable measures of both maternal and child behavior. But it introduces the major problem of the mother's ability, or desire, to present a true picture of her actual behavior. Errors may arise from involuntary misconceptions, faulty perceptions, or inaccurate recall.

Moreover, the mother may feel anxious or insecure about her normal behavior and may be unwilling to report it fully or without modifications that seem to her more acceptable to society. This is a more serious problem in establishing the validity of verbal reports as a measure of actual events because there are not sufficient cues to allow either discrimination of the mothers who consciously report inaccuracies or differentiation of the areas in which such errors occur.

Observation of behavior, when properly controlled, shows consistency of behavior and can give data on actual behavior under specific conditions. But it also involves the problems of the relationship between controlled conditions and the natural home environment and of the reliability of the sampling of maternal behavior. While both of these can be minimized by various techniques, they can be neither fully estimated nor controlled.

For the present research, dependency shown by the child at home and during the observation session was used as the constant child variable because earlier studies have shown certain relationships between specific maternal behavior and dependency (2, 3, 5, 6, 7). Maternal feelings about and behavior toward displays of dependency are the focus of the study. Replication of the findings of earlier studies will contribute to answering the two questions of primary concern here: (*a*) the relative adequacy and reliability of each method for studying maternal behavior; (*b*) the degree of similarity between reported and observed maternal behavior.

METHOD

Subjects

The *S*s were 30 mothers whose children, 18 girls and 12 boys, were enrolled in the Harvard Nursery School. The children's ages ranged from 3-0 to 4-1, with a mean age of 3-6. The mothers represented a highly selected group. Only four mothers had no college education; twenty-one had completed two or more years of college; five had completed one or more years of graduate work. It was expected, and found, that these mothers frequently would refer to publications on child rearing in order to explain their reasons for using a particular technique.

Procedure

The mothers were first observed for a 45-min. period while interacting with their children. Each mother was told that the *E* was interested in studying the child's play when she was present. She was reassured that no specific behavior was required and that she was free to do whatever she wished during the period. Both mother and child were then shown into the observation room, which was equipped with one-way vision windows, and the observer withdrew. After 30 minutes, the observer had the mother fill out a demographic questionnaire sheet in order to provide an experimental measure of her behavior toward the child's dependency solicitations when she was busy. As the observer withdrew again, the mother was told that there were 15 minutes remaining in the session.

At the end of the 45 minutes, *E* announced the end of this session and made arrangements for the interview. The interviews were conducted in the home by seven interviewers within two to four weeks after the observation session, and with only the mother present. Four interviews followed at longer intervals because of serious illness of mothers or children.

The training for the interviewers consisted of one

[1] This article is based on the author's Ph.D. dissertation completed at Radcliffe College in 1953 under the chairmanship of R. R. Sears, whose help is gratefully acknowledged.

month of intensive study of the interview schedule and the rating scales, as well as practice in actual interviewing. The practice interviews were criticized with reference to such factors as rapport with the mother, manner of interviewing, and adequacy of the answers when rated. The first few field interviews were also discussed with the interviewer. All interviews were recorded by means of an Audograph and scored directly from the records.

Equipment

The observation room contained a large table on which were several current magazines of various types, an upholstered chair, a child's table with two chairs, and a toy chest which could also be used as a bench or table. The toys were: (*a*) a set of finger paints and paper; (*b*) a set of small tools with thumbtacks and pieces of wood; (*c*) a set of colored blocks; (*d*) a jigsaw puzzle of a map of the United States; (*e*) a blackboard with erasers and chalk; (*f*) a partially dressed doll with the rest of its clothes; and (*g*) a set of interlocking blocks for building.

Scoring

To record the observed actions, a unit of behavior was defined as the equivalent of a simple sentence, although one element might be implied. The behavior of both the mother and the child was recorded on a scoring sheet containing a column for each category. The units were set down in rows in order of occurrence. Thus it was possible to indicate the sequence of interactions as well as the number and nature of elements scored for each person.

The child's dependency behavior was considered in terms of his stimulus and response activities; i.e., those actions which seemed to be primarily directed toward stimulating the mother to give help or attention which she did not volunteer, and actions that were primarily in response to situations which the mother seemed to set up. It was important to differentiate the conditions under which he sought help, the methods he used to solicit aid, and the kind of help or attention he desired. The mother's behavior was also thought of in terms of her stimulus and response actions, the latter being defined as behavior that resulted directly from specific stimulation by the child and served to end an on-going sequence. Although these actions might in turn stimulate further activity, they were differentiated on theoretical grounds from such maternal stimulus acts as degree of control exerted, specificity of control, and degree of contact imposed by the mother.

This conceptualization of the stimulus and response aspects of both child and mother behavior was used in the development of the categories for scoring the observed behavior. Table 1 presents the categories used for scoring both mother and child behavior. A category of positive social interaction was used to record verbal interchanges that were not concerned with dependency interactions, and therefore was given lowest priority in scoring and was not used in the final analysis of data. The observation session was timed and the end of each 15 minutes indicated on the scoring sheet.

Reliability was established with one other observer. Category reliability for those used varied from .75 to 1.00. Observer reliability for five sessions scored independently by the two observers ranged between .88 and .96, with a mean reliability of .94.

TABLE 1

SCORING CATEGORIES FOR OBSERVED BEHAVIOR

1. X—behavior in which the child engages to evoke the attention or affection of the mother; e.g., whining, leaning against her
2. Asking verbal help—asking for generalized directions or assistance; seeking information rather than direct aid
3. Asking material help—asking for specific physical help
4. Asking reward—asking for direct praise, reassurance, affection, or attention
5. Teaching—giving the child new information which is related to, but not dictated by, the situation; e.g., screwdrivers can be used to tighten screws or to lift out thumbtacks
6. Structurizing—giving generalized information from which the child may select his own behavior. The mother sets up the situation for the child, but gives him autonomy within it
7. Giving positive directions—giving specific directions or demands which are precise requirements for the child to meet
8. Giving negative directions—giving specific directions or demands which prohibit the child's ongoing behavior, but which do not offer alternatives to it
9. Doing—physically carrying out of tasks by the mother for the child, usually preceded by some directions; an attempt to demonstrate to the child how the task should be done
10. 0—asking permission to help when the child has expressed no need for aid
11. Giving reward—giving attention, nurturance, praise, or affection at the end of an action sequence or in response to the child's request
12. Comply—performance of a task which has been presented in the form of a demand, or verbal compliance to such a stimulus as the child's asking permission
13. Noncomply—lack of response to requested performance; including ignoring the demand
14. Punishment—verbal reprimands, physical punishment, disagreement, or derogation; usually comes at the end of an action sequence or serves to end a specific sequence
15. Leaves field—either participant indulges in activity or apparent thought which is unrelated to the other person, or to the total situation
16. Total observed dependency—summation of child scores under the first four categories; computed at the end of each session

The interview consisted of 36 open-ended questions relating to such variables as infant care and training, present demands made upon the child, amount and kinds of attention requested by the child at home, and the ways in which the mother responded to dependent behavior in the child. The scales on which ratings of mother behavior were made were grouped into the classifications of factual information, such as separations from the child, usage of "do" and "don't," and offering of help; information on feelings of the mother for the child, such as degree of warmth, nature of affectional relationship, and rejection by the mother; and information about the restrictions imposed upon the child, such as scheduling or feeding, restrictiveness of physical mobility, and overprotectiveness defined as

excesive protection of the child in a majority of situations where the child might be considered capable of independent action.[3]

Scales on child behavior were classified according to the nature of dependency solicitations (emotional, such as clinging or whining, or physical, such as wanting help while dressing), the conditions under which dependency occurs, and the areas in which the child strives for independence.

Interview scales which had a reliability of .60 or less were discarded and not used in analyzing the data. Since the remaining scales were divided into five points with 1 and 5 representing the extremes, and 3 the midpoint, two measures of rater reliability were used. The total rater reliability of the *E* and five other raters for nine interviews was .71 for complete agreement and .94 when one-point disagreements were included.

RESULTS

In comparing relative adequacy of these two methods, three measures were used: (*a*) consistency of observed behavior when a required task was interpolated; (*b*) replication of earlier findings on the antecedents of dependency behavior; and (*c*) similarity between observed behavior and reported behavior under similar conditions.

Table 2 presents the changes in behavior from the first 15 minutes of the session to the second 15 minutes, and from the first 15 minutes to the last. These are measured by the *t* test for the significance of differences between correlated means, and by product-moment correlations between the frequencies of each category for individual variability.

Both *structurizing* and *giving reward* decreased from the first to the second 15 minutes, possibly as a result of the child's increasing ease in the situation or of the mother's decreasing feeling of any need to stimulate the child to new and better activities.

From the first 15 minutes to the last 15 minutes, there was a significant decrease in *teaching* and *structurizing*, along with a significant increase in *give negative directions* and in *noncompliance*. Thus, when busy, the mothers became less rewarding and more restrictive. In general, however, there is evidence that individual mothers tended to be consistent in their behavior.

[3] A 5-page table giving the interview questions, and a 15-page table giving the interview rating scales have been deposited with the American Documentation Institute. Order Document No. 5635 from ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C., remitting $2.00 for microfilm or $3.75 for photocopies. Make checks payable to Chief, Photoduplication Service, Library of Congress.

TABLE 2

CHANGES IN MOTHER BEHAVIOR DURING THREE PARTS OF OBSERVATION SESSION

($N = 30$)

Categories	Mean Frequencies			t_{12}	t_{13}	r_{12}	r_{13}
	First 15 min.	Second 15 min.	Third 15 min.				
Teach	8.07	8.67	4.17	0.49	3.82***	.39*	.37*
Structurize	12.20	9.43	6.43	2.98***	6.14***	.64**	.36*
Give positive directions	23.53	24.27	22.07	0.43	0.85	.15	−.02
Give negative directions	3.10	3.83	4.23	1.43	2.02	.17	.38*
Give reward	13.67	10.23	15.23	2.92***	0.97	.30	.25
Comply	9.17	8.70	10.00	0.46	0.73	.48**	.21
Noncomply	3.11	3.67	14.67	0.07	7.71***	.72**	.21

* $p \leq .05$.
** $p \leq .01$.
*** $p \leq .001$.

Table 3 shows the correlations between five interview scales (selected because of their theoretical importance in the areas of warmth, frustration, and current restriction) and measures of dependency. Table 4 presents the correlations, in terms of *r*, between observed mother behavior and selected measures of dependency.

From Tables 3 and 4 the following significant relationships were found:

a. From interview data: *total dependency* is positively related to overprotectiveness; and *dependency* is positively related to rejection, especially in girls.

b. From observation data: *dependency* is negatively related to mother compliance and to mother giving reward, both of which are measures of lack of frustration; the more the mother *leaves the field* or *punishes* the child, the more frequently the child asks for material help, apparently to bring about an amicable close interaction; and the more the mother *complies* with the child's requests, the less verbal help or attention is requested.

c. From both observation and interview data: *dependency* in girls is positively related to warmth of the mother; and *dependency* is negatively related to amount of punishment given by the mother.

These replicate earlier results found on the relationship between mother behavior and dependency in children.

Since ratings on the techniques preferred by the mothers in their interviews followed the

TABLE 3
CORRELATIONS BETWEEN INTERVIEW ANTECEDENTS AND MEASURES OF DEPENDENCY

Interview Antecedents	Total Dependency on Interview	Total Dependency on Interview		Total Dependency on Observation	
		Boys	Girls	Boys	Girls
Schedule feeding	−.13	−.13	−.08	.71**	.04
Severity of weaning	.08	.13	.04	.14	.24
Rejection	.51**	.27	.65**	.10	.18
Affectional relations	.27	−.06	.47	−.01	.46*
Overprotection	.49**	−.03	.11	.03	.28

* Significant at .05 level.
** Significant at .01 level.

TABLE 4
CORRELATIONS BETWEEN OBSERVATION ANTECEDENTS AND MEASURES OF DEPENDENCY

Observation Antecedents	Total Dependency on Interview	Total Dependency on Observation	Ask Material Help	Ask Verbal Help
Comply	.18	−.36*	.09	−.36*
Noncomply	.07	.02	−.29	.22
Give reward	.07	−.36*	−.18	−.16
Leave field	−.20	.29	.51**	−.14
Punish	−.39*	.10	.40*	−.18

* Significant at .05 level.
** Significant at .01 level.

definitions set up for categories of observed behavior, reports of behavior when mothers were busy were compared with their behavior during the last 15 minutes of the observation session by means of the *t* test. No significant results were found. Using the same test for the entire observation session and reported techniques, only two significant differences were found. The mean observed use of structurizing was higher, at the .01 level, for those mothers who reported its use than for those mothers rated as not using it; giving positive directions was higher at the .05 level.

It was found that 70% of the mothers reported using techniques which they were also observed to use. Thus, by the sign test, ratings on techniques from the interview were positively related to the total observed techniques at the .05 level of significance. The *t* test revealed that mothers reported more accurately their use of "don't" than of "do." The mean observed use of negative directions throughout the session was higher, at the .05 level, for mothers who reported high usage of "don't" than for those who reported high usage of "do." In contrast, the mean observed use of positive directions was not significantly different for those mothers who reported high usage of "do" than for those who reported high usage of "don't." This result seems to indicate that mothers were more conscious of actions that they considered "bad" and were therefore able to make better estimates of how frequently they occurred.

DISCUSSION

The general consistency of behavior during the first half-hour of observation seems to indicate that the mothers were giving a reliable sampling of their normal behavior under similar conditions at home. When this is contrasted with the increased individual variability of behavior during the last 15 minutes, it can be hypothesized that the introduction of the task not only changed the mothers' activities, but also changed their set within the situation. Thus, mothers who approached the situation with a definite set as to how they should, or would, act throughout the 45 minutes, could not maintain this behavior when the external conditions were changed. But this alone does not invalidate the reliability of the method.

To supplement this measure, it was assumed that the method which best replicated previous findings on the relationship between mother behavior and dependency in the child would be the more adequate. Since the Iowa study (5) replicated findings of earlier studies, it was used as the criterion in this area. Mother behavior from the interview data seems to contribute more of the significant findings that replicate the Iowa study. But when *t*'s were computed on means of mother and child behavior during the last 15 minutes, it was found that increase in mother *noncomply* was positively related to increase in child *ask reward* at the .05 level, and mother *give reward* was negatively related to whining behavior at the .02 level. This relationship between warmth and whining is also similar to the Iowa report's finding on negative attention. Thus, the general results of this study replicated earlier findings on dependency, but neither method contributed more than the other.

A basic assumption for the study was that the observed behavior represented a criterion against which reported behavior could be validated. To test this, the mean and *SD* for changes in each observation category from the first to the last 15 minutes of the session were

determined. The mothers whose scores on each category lay more than one *SD* from the mean were identified. Most of the mothers' scores lay outside this range for one or another category, but only six mothers showed such a deviation on 40% or more of the categories.

The introduction of a new task was constant; therefore, it was reasoned that those mothers who showed a change of behavior greater than that shown by four-fifths of the mothers were probably giving a less representative sample of their usual behavior during the first half-hour of observation than were the majority. Whether these six differed because of the degree and kind of motivation under which they took part in the study, or because they were more anxious and thus more defensive about their normal behavior was explored by testing the hypothesis that their answers on the interview would show higher conformity to a group mean, especially for questions which might be anxiety-arousing, than the other mothers' answers.

Using the *t* test, this group was compared with the 24 other mothers on 10 interview questions, 5 of which were judged to involve some anxiety and 5, selected at random, which were considered to be neutral. There were no significant differences found. When the means and *SD*s for these 10 scales were computed, it was found that the 6 mothers were consistently within one sigma of the mean; of the other 24 cases, from 8 to 12 mothers on each scale fell outside this range. It was not, however, a consistent group who did so. This seems supportive evidence that "defensive" mothers can, and do, give interview information which may be misleading in its very normality. In addition, they may not be discovered during a brief observation under constant conditions.

From the preceding results, there is little evidence for deciding that either interviews or controlled observations are better for securing valid measures of actual mother behavior. Since observed behavior is consistent for most mothers; since both measures yield results that replicate previous findings on the antecedents of dependency; and since most mothers report similar behavior to that observed, the interview seems to be the preferable method because it permits study of a wider range of behavior in a shorter time than can an observation session. It is important to remember, however, that "defensive" mothers can be more easily discriminated by the variability of their behavior under observation conditions than by their answers to interview questions.

SUMMARY

In order to compare the adequacy of interviews and observations as methods, 30 mothers were studied by means of individual interviews and controlled observation while interacting with their children. Neither method was found to be completely superior to the other by the measures used: consistency of observed behavior when a required task was interpolated, replication of earlier findings on the antecedents of dependency behavior, and similarity between observed and reported behavior. In general, it was concluded that the interview was a preferable method only because it allows coverage of a wider range of behavior than does an observation. However, it was found that observation of behavior discriminated "defensive" mothers better than did the interview.

REFERENCES

1. CHAMPNEY, H. The variables of parent behavior. *J. abnorm. soc. Psychol.*, 1941, **36**, 525–542.
2. GEWIRTZ, J. L. Succorance in young children. Unpublished doctoral dissertation, State Univer. of Iowa, 1948.
3. LEVY, D. M. *Maternal overprotection.* New York: Columbia Univer. Press, 1943.
4. MERRILL, BARBARA. A measurement of mother-child interaction. *J. abnorm. soc. Psychol.*, 1946, **41**, 37–49.
5. SEARS, R. R., WHITING, J. W. M., NOWLIS, V., & SEARS, P. S. Child-rearing antecedents of aggression and dependency in young children. *Genet. Psychol. Monogr.*, 1953, **47**, 135–236.
6. SMITH, H. T. A comparison of interview and observation measures of mother behavior. Unpublished doctoral dissertation, Radcliffe College, 1953.
7. SYMONDS, P. W. *The psychology of parent-child relationships.* New York: D. Appleton-Century, 1939.

Received July 29, 1957.

NEURAL SATIATION, REACTIVE INHIBITION, AND INTROVERSION-EXTRAVERSION[1]

ALLAN RECHTSCHAFFEN[2]

Northwestern University

EXPANDING upon earlier theorizing by Pavlov (19), Eysenck (6) has hypothesized that extraverts develop more reactive inhibition (I_R) than introverts. Eysenck further noted similarities between I_R and neural satiation, the construct proposed by Köhler and Wallach (18) to underlie figural aftereffects. The similarities of the constructs and the phenomena from which they were inferred were also noted by Köhler and Fishback (17) and Duncan (4). On the basis of such similarities, Eysenck further proposed that extraverts are more susceptible to the development of neural satiation than introverts.

In his first test of these hypotheses, Eysenck (6) compared the kinesthetic aftereffects of introverts and extraverts. The introvert and extravert groups were defined by the dual criteria of cutoff scores on Guilford's R (Rhathymia) Scale (8) and psychiatric diagnoses; obsessionals, reactive depressives, and anxiety neurotics were included in the introvert group, while hysterics and psychopaths were included in the extravert group. The results were interpreted as showing significantly greater aftereffects for extraverts, hence supporting the hypothesis. In a more recent study, Eysenck (5) correlated reminiscence scores (which he proposed as a measure of I_R) on a pursuit rotor task with the extraversion-intraversion scores of university students on the Maudsley Personality Inventory, again obtaining results that he interpreted as confirming his hypothesis. Eysenck has also cited the work of Franks (7) which showed that introverts developed conditioned responses more readily than extraverts as providing further support for his theories. While some of the above findings certainly appear suggestive, certain procedural and statistical aspects of these investigations cast some doubt on Eysenck's interpretation. These matters are discussed at a later point.

The present study was designed to investigate further the hypotheses cited above. Three hypotheses were investigated: 1. Extraverts develop greater visual aftereffects than introverts. This hypothesis follows directly from Eysenck's theory, if one can assume that satiability is a fairly general trait extending across modalities. 2. Extraverts develop more I_R on a motor-learning task than introverts. In the present investigation, inverted alphabet printing was used to measure I_R. This choice of task seems appropriate in that several investigators (e.g., 2, 20) have demonstrated large amounts of I_R to build up with massed practice in inverted alphabet printing. 3. Amounts of aftereffect and I_R exhibited by individuals are positively correlated. This hypothesis follows from the thinking of Eysenck, Köhler and Fishback, and Duncan regarding the similarities of the two phenomena. In his summary of the literature on this topic, Duncan has pointed out that not only are I_R and aftereffects grossly similar in that they both represent response alteration following continued response evocation, but that they also have strikingly similar accumulation and decay functions. Such similarities lead to the expectation that individual differences in I_R would be correlated with differences in aftereffect, assuming either that the same neural substrate is involved in both phenomena or that individual susceptibility to the phenomena is relatively independent of the particular neural area being stimulated.

METHOD

The over-all plan of the investigation involved obtaining from each *S* a measure of introversion-extraversion, I_R, and visual aftereffect.

Subjects and Measurement of Introversion-Extraversion

Guilford's R Scale, which had been administered in group form to an entire introductory psychology class, was used to define the introversion-extraversion dimension, high R scores representing extraversion. From the class, 96 *S*s were obtained as volunteers, a few *S*s being

[1] This paper is based upon a doctoral dissertation submitted to Northwestern University. The author is grateful to Janet A. Taylor for her guidance in this research.

[2] Now at the University of Chicago Medical School.

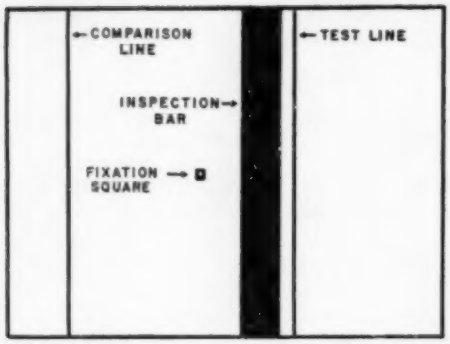

FIG. 1. RELATIONSHIP OF TEST AND INSPECTION FIGURES IN THE VIEWING AREA
(Figures are superimposed so that their fixation squares coincide)

requested to participate in order to obtain an equal number with R scores falling above and below the class mean and an equal number of males and females. The mean R score of the 96 *Ss* was 38.8; the *SD* was 11.5; the range was 8–70; the distribution was fairly symmetrical, tending towards bimodality.

Procedure

Ss were tested individually on the inverted alphabet-printing and visual aftereffect tasks during a single experimental session lasting approximately 45 min. The order of task presentation was counterbalanced, half of the *Ss* being given the alphabet printing first and half the visual aftereffect task first. Within each of these groups, the same counterbalancing was maintained with regard to sex and position above or below the sample mean on R scores. For all *Ss*, a rest of 5 min. separated the administration of the two tasks.

Inverted Alphabet Printing

Ss were told that the inverted alphabet-printing task was designed to measure a facet of their intelligence. They were given standard instructions for the task (e.g., 2), including a demonstration that seven letters appeared the same inverted as they did right side up. *Ss* worked in 20-page booklets, each page containing a 10 × 6 matrix of ½-in. squares.

Ss were first given five 30-sec. trials separated by 30-sec. rest intervals (distributed practice). At the end of each 30-sec. practice interval, *Ss* were instructed to turn to the next page of the test booklet, sit quietly for the 30-sec. rest interval, and to start on the new page with the next consecutive letter (following the last one they had written on the preceding page) when *E* gave the signal.

Following the distributed practice, *Ss* were given a 1-min. rest interval and then 10 consecutive 30-sec. trials (5 min. of massed practice). At the end of each 30-sec. trial, *E* signaled to *S* to turn to the next page of the test booklet and proceed immediately with the next consecutive letter.

After 49 *Ss* had been run under the conditions described above, it was decided to include a reminiscence measure in the study. Hence, the remaining 47 *Ss* were given a 1-min. rest following the last trial of massed practice and two consecutive 30-sec. trials following the rest. Of the 47 *Ss* thus tested, 25 *Ss* had R scores of 39 or above, and 22 *Ss* had R scores of 38 or below.

Visual Aftereffect

A figure (see Fig. 1) similar to one employed by Wertheimer (21) was used.[1] During inspection, only the fixation square and inspection bar were shown, theoretically satiating the region corresponding to the visual space to the right of the fixation square. When the test figure (fixation square, test line, and comparison line) were subsequently presented, the satiation should have become manifest in an apparent displacement of the test line away from the fixation square. By having *Ss* adjust the comparison line to appear the same distance from the fixation square as the test line, these apparent displacements could be measured as differences between preinspection and postinspection settings of the comparison line.

An apparatus similar in principle to one employed by Hammer (9) was used. The essential components were two clear-glass plates firmly set in a slotted frame. On one plate were fixed the fixation square and inspection bar of the inspection figure. By appropriate movement of this plate in the slotted frame, the test and inspection figures could be alternately exposed in the opening of a reduction screen. On the other glass plate, which lay proximal to *S* and almost immediately next to the first one, was fixed the comparison line. By turning an adjustment wheel which was attached to a pulley system on the frame, *Ss* could move the plate containing the comparison line either left or right. A hairline indicator attached to the comparison-line plate moved against a millimeter rule located at the rear of the apparatus, thus enabling *E* to make readings of *Ss*' settings of the comparison line. A translucent milk-glass plate was situated behind (distal to *S*) the two plates cited above. The viewing area was illuminated from behind this translucent plate, providing a homogeneous field, except for the test and inspection figures, which was free from distinguishing cues.

After being instructed in setting the comparison line the same distance from the fixation square as the test line, *Ss* were given three practice trials of four adjustments each. (The terminology of referring to a set of four adjustments as a trial will be maintained throughout this report.) *Ss* were trained to make the four adjustments within a trial in rapid succession, and most *Ss* learned to make four adjustments in about 30 sec. The first and third adjustments of a trial were always ascending, i.e., the comparison line was set too close to the fixation square by *E* and had to be moved further away by *S*. The second and fourth adjustments were descending. The exact presettings of the comparison line were varied by *E* in an unsys-

[1] A more exact description of the figure and the apparatus used, including all the dimensions involved in each, can be obtained from the original dissertation which is available from University Microfilms, Mic 56-5336, for $1.50.

tematic manner. On all adjustments, *S*s were allowed to move the comparison line in either direction in order to make finer adjustments.

Following practice trials, *S*s were given a 2-min. rest during which they were given instructions for the experiment proper. The instructions strongly stressed the importance of maintaining constant, faithful fixation on the fixation square during inspection and test periods. Then the following schedule was maintained: *S*s were given a preinspection trial, a 1-min. rest, 40 sec. of fixation on the inspection figure, and a postinspection trial. Two additional determinations of aftereffect were made for each *S*, making a total of three sequences of preinspection trial, rest, inspection, and postinspection trial. A rest of 3 min. was allowed between each sequence.

All adjustments of the comparison line were made monocularly with the right eye, while the inspection figure was always fixated with the left eye.[4] This procedure was introduced, because, during pilot work in which inspection and test were accomplished with the same eye, *S*s frequently reported strong negative afterimages of the inspection figure in postinspection trials. These afterimages usually caused *S*s much difficulty in maintaining fixation and making adjustments.

RESULTS

Visual Aftereffect

For all 96 *S*s, the mean amount of aftereffect (differences between preinspection and postinspection trials) was in the expected direction for all three inspection periods. Each of these means (4.1, 4.0, and 4.2 mm. respectively) was significantly different from zero ($p < .01$). Differences between the three preinspection trials were negligible and not significant, indicating that there was probably fairly complete dissipation of aftereffect during the 3-min. rest interval following the first and the second postinspection trials. This consideration enables us to view the three determinations of aftereffect as independent measures for purposes of estimating reliabilities. A single aftereffect score was determined for each *S* by summing his aftereffects for the three inspection periods. The reliability of these scores as estimated by Hoyt's r (10) was .52. Intercorrelations of the three independent measures were also computed. These were .25 for Trials 1 and 2, .25 for Trials 1 and 3, and .27 for Trials 2 and 3 ($p < .02$ in each case).

Thus the task did produce systematic, significant aftereffects as would be predicted by satiation theory, and significant, albeit low, reliability was found among the three measures. However, the results of the study do not support the hypothesis that there is a relationship between amount of visual aftereffect and R scores, our personality measure. The r between the aftereffect scores and the R scores for all 96 *S*s was .14. While the relationship was in the predicted direction, the magnitude of the correlation fell short of the .20 required for significance at the .05 level of confidence with 94 *df*.

In view of the possibility that only extreme introversion and extraversion are related to amount of aftereffect, an additional analysis was undertaken. Two subgroups were formed from the 96 *S*s. These groups were the "extreme extraverts," which consisted of the 35 *S*s with the highest R scores (44 or above) and the "extreme introverts," which consisted of the 35 *S*s with the lowest R scores (32 or below). The aftereffects of the 35 extreme extraverts were compared with the aftereffects of the 35 extreme introverts. As can be seen graphically in Fig. 2, the extreme extraverts appear to have somewhat larger aftereffects. However, the difference was not statistically significant. The extreme extraverts had a mean aftereffect score of 13.6 mm. as compared to 10.3 mm. for the extreme introverts, the t value for the difference being only 1.07.

Inspection of Fig. 2 also reveals in each group the expected development of aftereffect with inspection and dissipation with rest. It will be noted that the extreme extraverts tend to have lower preinspection settings than the extreme introverts. In none of the three trials, however, did the preinspection settings of the two groups differ significantly, nor were preinspection settings for all 96 *S*s significantly correlated with amount of aftereffect.

Inverted Alphabet Printing

For an I_R measure, an "inhibition score" was calculated for each *S* by subtracting the number of letters printed (correctly or incorrectly) during the last five trials of massed

[4] It is possible that the magnitude of visual aftereffects is reduced by using different eyes for inspection and test periods, but the author knows of no systematic studies on the subject. One possible interpretation of the bilateral transfer of satiation effects in the present study is that since there is bilateral representation of the macula in the occipital lobes, since the macula subtends a visual angle of approximately 3°, and since the inspection bar was included within 1°40′ of visual angle from the center of the fixation square, inspection could have produced satiation in corresponding portions of both occipital lobes.

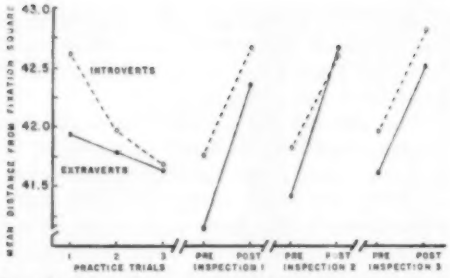

FIG. 2. MEAN DISTANCE PER ADJUSTMENT OF COMPARISON LINE FROM INSPECTION SQUARE IN MM. FOR PRACTICE, PREINSPECTION, AND POSTINSPECTION TRIALS, COMPARING 35 EXTREME EXTRAVERTS AND 35 EXTREME INTROVERTS. AMOUNT OF AFTEREFFECT IS SHOWN AS THE DIFFERENCE BETWEEN PREINSPECTION AND POSTINSPECTION TRIALS

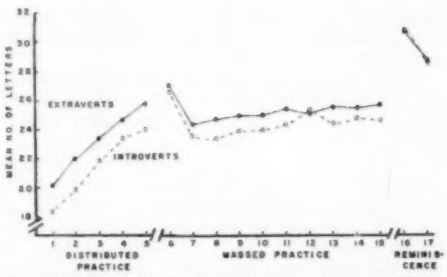

FIG. 3. COMPARISON OF PERFORMANCE OF 35 EXTREME EXTRAVERTS AND 35 EXTREME INTROVERTS ON INVERTED ALPHABET PRINTING FOR TRIALS 1–15[6]

practice from the number of letters printed during the five trials of distributed practice. Thus the higher the score the higher the I_R, e.g., a greater score during the early, distributed practice than during the later massed practice would indicate more I_R than a poorer score during distributed practice.

Using the inhibition scores thus derived, the results of the study do not support the hypothesis that there is a relationship between the accumulation of I_R and R scores. The r between the inhibition scores and the R scores for all 96 *S*s was .18.[5] While the relationship was in the predicted direction, it was not statistically significant.

A comparison was also made between the inhibition scores of the 35 extreme extraverts and 35 extreme introverts (see Fig. 3). It is readily apparent that although the performance curve of the extreme extraverts tends to remain above that of the extreme introverts, the slopes of the curves remain essentially similar throughout, contrary to the divergence expected if more I_R were being built up in one group than another. This finding would suggest that there is no substantial relationship between R score and accumulation of I_R. The extreme extraverts had a mean inhibition score of −10.9 as compared to a mean inhibition score of −15.7 for the extreme introverts. The difference was in the predicted direction but was not significant ($t = 1.26$). An analysis of variance for all 15 trials showed no significant difference between the groups on level of performance and no significant interaction between groups and trials.

It should be noted that the inhibition measure discussed above differs from that used by Eysenck (5), which was reminiscence. Since reminiscence represents a recovery from inhibition rather than inhibition per se, the use of the reminiscence score as an index of I_R appears to rest on the assumption that amount of recovery is positively related to the amount of I_R accumulated. While such an assumption seems reasonable, Eysenck (6) had postulated that not only do extraverts develop more I_R, but also that they dissipate it more slowly. Thus conflicting predictions result from Eysenck's theory when reminiscence scores are used. On the one hand, it seems reasonable to assume that if extraverts develop more I_R, they should show more recovery, i.e., a greater difference between performance at the end of massed trials and performance on reminiscence trials. On the other hand, since extraverts theoretically recover more slowly from I_R, it might be expected that they would show less reminiscence. Possibly, the two factors could both be operating and mask each other.

In spite of the above objection, a reminis-

[5] The r of .18 between inhibition scores and R scores is significant at the .05 level of confidence with 94 *df* when a *one-tailed* test is used as the criterion of significance. However, it should be noted that the occurrence of one such r, significant at the .05 level with a one-tailed test, from among the four r's calculated in testing the major hypotheses, is not very unlikely on the basis of chance alone.

[6] The reminiscence trials are for 20 extraverts and 17 introverts; they should not be used for comparing extraverts and introverts on absolute amount of reminiscence, since the preceding massed practice curves were based on a larger *N*.

cence score was obtained for 47 *S*s to offer some basis for comparison with Eysenck's study. Reminiscence scores were obtained by subtracting the number of letters printed during the last two trials of massed practice from the number of letters printed during the reminiscence trials.

The r between inhibition and reminiscence scores for the 47 *S*s receiving both measures was .35, which is significant at the .02 level of confidence. While the above correlation is probably somewhat spurious by reason of the fact that the number of letters printed on the last two trials of massed practice enters into both the inhibition and reminiscence scores, the finding tends to support the assumption that the more I_R accumulated, the greater the reminiscence. However, as was pointed out above, differential rates of recovery for extraverts and introverts may be confounded in the reminiscence measure with amount of I_R accumulated.

The r between the reminiscence scores and the R scores for the 47 *S*s who received reminiscence trials was .08. With 45 *df*, an r of .29 is required for significance at the .05 level of confidence. Of the 35 extreme extraverts, 20 received reminiscence trials, and of the 35 extreme introverts 17 received reminiscence trials. These 20 extreme extraverts and 17 extreme introverts achieved mean reminiscence scores of 8.8 and 8.6 respectively, the difference between these means yielding a t value of only .06.

Relationship Between Visual Aftereffect and I_R

Although the magnitude of neither I_R nor aftereffects was significantly related to introversion-extraversion scores, it is still possible that the magnitude of aftereffects is correlated with the magnitude of I_R as hypothesized earlier. The r between the aftereffect and inhibition scores for the 96 *S*s was, however, only .04. The r between the aftereffect and reminiscence scores was only $-.02$.[7]

[7] Since the original writing of this paper, the author has learned of a study (Meier, M. J. Interrelationships among personality variables, kinesthetic figural aftereffects and reminiscence in motor learning. Unpublished doctoral dissertation. Univer. of Wisconsin, 1956) which showed a significant *negative* correlation between amount of reminiscence in inverted alphabet printing and amount of kinesthetic aftereffect.

DISCUSSION

The results do not support the hypotheses that amount of aftereffect is related to introversion-extraversion, that accumulation of I_R is related to introversion-extraversion, or that the amount of aftereffect exhibited by an individual is related to his accumulation of I_R. While the results were in the expected direction in each case, the degree of relationship demonstrated was uniformly small, and in no case did it reach statistical significance.

There are several differences between the present study and those of Eysenck which may be partly responsible for the apparent divergence in results, e.g., different *S*s, differences in modalities being tested, task differences, and differences in experimental procedure. Since none of these differences appear to violate any of the boundary conditions specified by Eysenck for testing his hypotheses, however, it is clear his theories concerning the relationship of introversion-extraversion, I_R, and aftereffect are not supported under the experimental conditions described in this study. And, as previously mentioned, there are grounds for uncertainty concerning Eysenck's findings themselves.

Consider first Eysenck's study (6) relating kinesthetic aftereffects to introversion-extraversion. It will be remembered that the aftereffect scores in the present study were obtained from the differences between preinspection and postinspection settings of the comparison line on each trial. Eysenck's procedure was different. He obtained a single preinspection setting (an average of four adjustments) and four postinspection settings, one after each of four successive inspection periods. As aftereffect scores for each of the four trials, Eysenck used the per cent change from the original preinspection setting. Since he was also interested in relating rate of dissipation of aftereffect to personality, Eysenck also obtained settings after 5 min. of rest and again after an additional 10 min. of rest. The major evidence offered by Eysenck for the difference in aftereffect between introverts and extraverts was a Hotelling's T test which invalidated the null hypothesis at between the .01 and .05 levels of confidence. The calculation of Hotelling's T, an over-all test of significance, was based upon nine separate aftereffect scores which were as follows:

1. Average per cent change after 30 sec. of inspection; 2. average per cent change after 60 sec. of inspection; 3. average per cent change after 90 sec. of inspection; 4. average per cent change after 120 sec. of inspection; 5. sum of the above four scores; 6. maximum single per cent change obtained from each *S*; 7. average per cent change after 5 min. of rest; 8. average per cent change after an additional 10 min. of rest; 9. sum of the above two scores.

First, it appears likely that Scores 1–4 are not independent measures of aftereffect. In discussing his experimental procedure, Eysenck gave no indication that there were any rest pauses separating the four inspection periods, aside from the time taken to make the postinspection settings. Since aftereffects following comparable periods of inspection are known to last up to 90 sec. after inspection (9) and have been found to cumulate (21), we might expect that there would be an accumulation of undissipated aftereffect from trial to trial in Eysenck's procedure. While the amount of such accumulation could not be estimated from Eysenck's results (he used progressively longer inspection periods in the four successive trials, and accumulation of undissipated satiation is confounded with increased inspection time), it seems likely that some accumulation resulted and that the successive trials in his study cannot be considered to be independent measures of aftereffect. In addition, there is an obvious lack of independence among the first four scores and Scores 5 and 6, as well as among Scores 7, 8, and 9. It would appear to be a questionable procedure to use for an over-all test of significance such a series of nonindependent, spuriously intercorrelated scores, some of which do little more than superfluously repeat original trends.

Eysenck also computed *t* tests between introverts and extraverts for the nine scores considered individually. Of the nine *t*'s, one (for the 30-sec. inspection period) was barely significant at the .05 level of confidence using a one-tailed test. The other eight scores were significant at approximately the .10 level only, using a one-tailed test. While Eysenck's findings are suggestive, they are far from striking, and the failure to replicate his results in the present study is not particularly surprising.

A final question arises as to the relative adequacy of the aftereffect measures. The estimated reliability of the aftereffect scores in the present study was .52 as compared to estimated reliabilities of .78 and above reported by Eysenck for his measure of kinesthetic aftereffect. At face value, the measure used in the present study seems relatively poor for singling out a highly stable individual characteristic of satiability. Closer examination, however, discloses important differences in the derivation of the respective scores which makes the apparent relative unreliability of our own scores open to considerable question.

First, it will be noted that Eysenck's reliability coefficients were obtained by Hoyt's *r* which depends, in essence, upon the correlations among the several trials. As noted previously, however, the four test trials in Eysenck's measure of kinesthetic aftereffect are probably nonindependent by reason of the accumulation of undissipated aftereffect from trial to trial. For purposes of reliability estimation, this type of nonindependence must be considered as producing spuriously high correlations. Further, it will be recalled that Eysenck used a single preinspection setting as a base against which to calculate per cent change for all four postinspection trials. Insofar as the single preinspection setting enters into all four aftereffect measures, and insofar as we can expect interindividual variation in preinspection settings, the reliability of aftereffect is confounded with the constancy of the preinspection setting, once more producing a spuriously high reliability coefficient.

It is entirely likely that kinesthetic aftereffects are more reliable than visual aftereffects, because they are not so subject to the vagaries of fluctuating fixation. However, in view of the errors involved in Eysenck's estimation of reliability, this has not yet been adequately demonstrated. Certainly, the differences in reliability are not so great as uncritical evaluation of the reported reliabilities would lead us to believe.

Wertheimer (22) has reported a retest reliability of .82 for visual aftereffects. However, his measure of aftereffect was based on postinspection settings alone, rather than on differences between preinspection and postinspection settings. Inasmuch as we can expect interindividual variation in constant error (apart from the development of aftereffect), and inasmuch as we can expect some reliability of constant error in successive tests, Wertheimer confounds reliability of aftereffect with the reliability of constant error. For example, using Wertheimer's procedure with our own data,

the estimated reliability was raised from .52 to .92

Wertheimer himself has more recently called attention to the importance of considering constant errors in the measurement of aftereffects (23).

Now we turn to the evidence Eysenck has presented as demonstrating the hypothesized relationship between introversion-extraversion and I_R. Prior to performing his own experiment on this particular hypothesis, Eysenck (6) cited the work of Franks (7) as supporting his theory. Franks found that introverts (dysthymics) showed a higher level of performance in eyelid conditioning than extraverts (hysterics), and Eysenck held this difference to result from greater I_R for the extraverts. There is reason to question this interpretation. According to the Hullian definition (11), which Eysenck has followed, I_R is produced by response evocation, the magnitude being dependent on the amount of work involved. Further, I_R cumulates under conditions of massed practice and becomes manifest in performance decrement (when groups receiving massed practice are compared with groups receiving distributed practice). Franks presented no indication of the intertrial interval employed in his conditioning procedure nor of the amount of work involved in each response, although it should be noted that the typical intertrial intervals and the amount of work involved in eyelid conditioning are not favorable for the accumulation of I_R. Had the conditions produced accumulation of I_R, however, and had the differences between extraverts and introverts derived from differences in I_R, then it would be expected that the differences in performance between the groups would increase as conditioning progressed, the amount of I_R cumulated becoming relatively greater for the extraverts. Inspection of the acquisition curves obtained by Franks shows this not to be the case. The two curves differ on the initial conditioning trials and show no systematic tendency to diverge on subsequent trials, an apparent contradiction of the I_R hypothesis.

In his own experiment relating introversion-extraversion to I_R, Eysenck (5) employed the pursuit-rotor task, giving three 5-min. periods of massed practice separated by two 10-min. rest periods. As was discussed previously, Eysenck used reminiscence (following the two rest periods) as a measure of I_R. It is important to note that Eysenck used the difference between the last three trials of prerest practice and the first trial of postrest practice as his reminiscence measure. This procedure was in contrast to the measure advocated by Ammons (1) who argued that a correction be made for warm-up (the rapid initial rise in performance following the first post-rest trial). Ammons reasoned that reminiscence should be measured as the difference between the last trial of pre-rest practice and the asymptote of the warm-up increment, so that warm-up effects might not mask the full amount of reminiscence occurring. Eysenck, however, has followed the argument of Denny, Frisby, and Weaver (3) that what Ammons has called "warm-up" is actually the extinction of ${}_sI_R$. By this reasoning, the use of the reminiscence measure suggested by Ammons would confound the effects of I_R and ${}_sI_R$.

Eysenck's results showed an r of .29 between the first reminiscence score (following the first rest period) and an introversion-extraversion measure. For his N of 50 *Ss*, this was significant at the .02 level of confidence using a one-tailed test. The r between the second reminiscence score and the introversion-extraversion measure was only .10; that between both reminiscence scores and introversion-extraversion was .22; neither value is significant at the .05 level of confidence. Eysenck interpreted these results as showing a significant relationship between I_R and the personality measure. He explained that a lower correlation between introversion-extraversion and the second reminiscence measure was to be expected. Citing earlier findings by Kimble (12, 13, 14, 15, 16), Eysenck stated that ${}_sI_R$ can be expected to develop late in practice, and reasoned further that this increased amount of ${}_sI_R$ at the time of the second reminiscence measure probably increased error variance so as to make the second reminiscence measure a less reliable index of I_R than the first. Eysenck's case would appear to depend on the following conditions: there would have to be a fairly substantial difference in error variance between the two reminiscence situations to result in a shift from .29 to the .10 correlation; this in turn would have to result from very substantial differences in ${}_sI_R$ between the two reminiscence situations. Now Eysenck has suggested that the initial rise in

performance following rest is the result of extinction of ${}_{s}I_{R}$. On this basis, it would be expected that the "${}_{s}I_{R}$ extinction increment" following the second rest period would be substantially greater than that following the first rest period (assuming that amount of ${}_{s}I_{R}$ extinction is positively related to amount of ${}_{s}I_{R}$ accumulated during prerest practice). Inspection of the graph of performance presented by Eysenck, however, shows this not to be the case. The second "${}_{s}I_{R}$ extinction increment" is only barely greater than the first. In view of this internal inconsistency in Eysenck's interpretation and data, and in view of dangers generally inherent in selecting significant findings as more valid than insignificant findings on a speculative basis, we are once more led to conclude that Eysenck's findings do not support his hypotheses conclusively.

While in general it may be said that all studies thus far performed on the relationship between introversion-extraversion and aftereffects, and between introversion-extraversion and I_R, have yielded results in the direction predicted by Eysenck's theory, in no case has the degree of relationship been shown to be very strong, and rarely have the findings been statistically significant. Acceptance of the theory seems to await more conclusive supporting evidence.

Leaving the personality variable and considering the nonsupportive evidence concerning the hypothesized similarity of I_R and neural satiation, one must be cautious regarding the possible interpretations. Certainly, the results of the present study would indicate that the two phenomena, as they are measured here, are not correlated. However, this does not mean that the phenomena are dissimilar; on a descriptive level there is no denying their similarity. Furthermore, it is even possible that I_R and neural satiation are based on the same neurological processes. If there is no great intrasubject consistency in the satiability of different neural areas, then the negative results of the present investigation could derive from the fact that different neural areas were being tapped in the measurement of the two phenomena. For the same reason, it might not be surprising to find that separate measures of I_R determined on different tasks were not correlated. Response correlation is useful as supportive evidence for similarities between phenomena, but the method is by no means decisive for ruling out similarities in the underlying processes.

SUMMARY

Expanding upon earlier theorizing by Pavlov, Eysenck has hypothesized that extraverts would show greater figural aftereffects (or neural satiation) and more reactive inhibition (I_R) on a motor task than introverts. Several investigators had noted similarities between neural satiation and I_R and had hypothesized that the neural mechanisms underlying both phenomena might be the same.

These various hypotheses were investigated in the present study by testing 96 *S*s on introversion-extraversion, visual aftereffects, and I_R. Introversion-extraversion was defined in terms of scores on Guilford's Rhathymia Scale. Visual aftereffects were measured with a specially constructed apparatus, and two measures of I_R were obtained from an inverted alphabet-printing task.

The results showed that the introversion-extraversion (Rhathymia) scores were not significantly correlated with either amount of visual aftereffect or I_R measures. Further, amount of visual aftereffect was not significantly correlated with the I_R measures. An additional analysis was made comparing the 35 *S*s with the highest Rhathymia scores (extraverts) with the 35 *S*s with the lowest Rhathymia scores (introverts). The two groups also did not differ significantly on the aftereffect and I_R measures.

The results were interpreted as failing to support the hypotheses concerning the relatedness of the three variables, and the findings were compared to those of previous investigations.

REFERENCES

1. AMMONS, R. B. Acquisition of motor skills: I. Quantitative analysis and theoretical formulation. *Psychol. Rev.*, 1947, **54**, 263–281.
2. ARCHER, E. J. Postrest performance in motor learning as a function of prerest degree of distribution of practice. *J. exp. Psychol.*, 1954, **47**, 47–51.
3. DENNY, M. R., FRISBY, N., & WEAVER, J. Rotary pursuit performance under alternate conditions of distributed and massed practice. *J. exp. Psychol.*, 1955, **49**, 48–54.
4. DUNCAN, C. P. On the similarity between reactive inhibition and neural satiation. *Amer. J. Psychol.*, 1956, **69**, 227–235.

5. EYSENCK, H. J. Reminiscence, drive, and personality theory. *J. abnorm. soc. Psychol.*, 1956, **53**, 328–333.
6. EYSENCK, H. J. Cortical inhibition, figural after-effect, and theory of personality. *J. abnorm. soc. Psychol.*, 1955, **51**, 94–106.
7. FRANKS, C. M. Conditioning and personality: A study of normal and neurotic subjects. *J. abnorm. soc. Psychol.*, 1956, **2**, 143–149.
8. GUILFORD, J. P. *An inventory of factors STDCR.* Beverly Hills, Calif.: Sheridan Supply, 1942.
9. HAMMER, ELAINE R. Temporal factors in figural after-effects. *Amer. J. Psychol.*, 1949, **62**, 337–354.
10. HOYT, C. J. Test reliability estimated by analysis of variance. *Psychometrika*, 1941, **6**, 153–160.
11. HULL, C. L. *Principles of behavior.* New York: Appleton-Century, 1943.
12. KIMBLE, G. A. An experimental test of a two-factor theory of inhibition. *J. exp. Psychol.*, 1949, **39**, 15–23.
13. KIMBLE, G. A. Performance and reminiscence in motor learning as a function of the degree of distribution of practice. *J. exp. Psychol.*, 1949, **39**, 500–510.
14. KIMBLE, G. A., & BILODEAU, G. A. Work and rest as variables in cyclical motor learning. *J. exp. Psychol.*, 1949, **39**, 150–157.
15. KIMBLE, G. A., & HORENSTEIN, B. R. Reminiscence in motor learning as a function of length of interpolated rest. *J. exp. Psychol.*, 1948, **38**, 239–244.
16. KIMBLE, G. A., & SHATEL, R. B. The relationship between two kinds of inhibition and the amount of practice. *J. exp. Psychol.*, 1952, **44**, 355–359.
17. KÖHLER, W., & FISHBACK, J. The destruction of the Müller-Lyer illusion in repeated trials: II. Satiation patterns and memory traces. *J. exp. Psychol.*, 1950, **40**, 398–410.
18. KÖHLER, W., & WALLACH, H. Figural after-effects: An investigation of visual processes. *Proc. Amer. Phil. Soc.*, 1944, **88**, 269–357.
19. PAVLOV, I. P. *Lectures on conditioned reflexes.* Vol. II. London: Lawrence & Wishart, 1941.
20. SCHUCKER, R. E., STEVENS, L. B., & ELLIS, D. S. A retest for conditioned inhibition in the alphabet-printing task. *J. exp. Psychol.*, 1953, **46**, 97–102.
21. WERTHEIMER, M. Figural after-effects in schizophrenic and normal subjects. *Amer. Psychologist*, 1954, **9**, 492. (Abstract)
22. WERTHEIMER, M. The differential satiability of schizophrenic and normal subjects: A test of a deduction from the theory of figural after-effects. *J. gen. Psychol.*, 1954, **51**, 291–299.
23. WERTHEIMER, M. Constant errors in the measurement of figural after-effects. *Amer. J. Psychol.*, 1954, **67**, 543–550.

Received July 31, 1957.

BODY IMAGE AND ASYMMETRY OF BODY REACTIVITY[1]

SEYMOUR FISHER

Baylor University College of Medicine

IN STUDIES (6, 7, 8) focused on the manner in which the individual experiences his body boundaries, it was shown that persons differ greatly in the degree to which they conceive of their body boundaries as definite and firm as contrasted to indefinite and weak. One of the most meaningful of the observed correlates of degree of boundary definiteness had to do with site of psychosomatic symptomatology. The data indicated that persons who develop psychosomatic symptoms in the exterior body layers (e.g., neurodermatitis) are characterized by definite body-image boundaries, whereas individuals with body interior symptoms (e.g., stomach ulcers or spastic colitis) are marked by poorly defined body-image boundaries. Further, degree of boundary definiteness appears to play a determining part in this relationship and is not simply a reflection of the experiencing of pain and discomfort at particular body sites.

When it was found that body-image boundary definiteness was linked with body symptomatology, it followed that it may also be linked with more basic physiological variables. If one assumes that the development of psychosomatic symptoms at a given site indicates generally an exaggerated physiological reactivity or sensitivity at that site, one may hypothesize that prior to the development of obvious clinical pathology, individuals with definite boundaries are likely to manifest a higher degree of reactivity in their exterior body layers (skin, muscle, and peripheral blood vessels) and a lower degree of body interior reactivity than persons with indefinite boundaries. It was predicted that people with definite boundaries typically yield a gradient of physiological reactivity such that the body exterior is relatively more responsive than the body interior. Persons with indefinite boundaries were hypothesized to show a reversal of this gradient such that greatest reactivity to stimuli occurs in the body interior and relatively less in the exterior body layers. Substantiation for this hypothesis was obtained in two studies (7, 8) relating measures of boundary definiteness to measures of physiological responses at different body levels (e.g., MAP, GSR, heart rate).

This relationship of exterior-interior gradients of reactivity and an index of body image led to the proposition that one's body may be conceptualized as a screen upon which body-image attitudes are projected and that the pattern of this projection may influence the pattern of physiological reactivity. Thus, in applying this proposition to the body-image boundary dimension, one may hypothesize that the more definite and firm an individual's concept of his body boundary, the more likely he is to build up organized tonus sets and systems involving readiness to respond in the exterior body layers. In an oversimplified way, this process may be compared to the high tonus level that develops in an arm muscle when one arrives at a definite conclusion about a course of action that involves an arm movement and is about to initiate that movement. One's long term attitudes toward various sectors of one's body may result in selective modes of response in those sectors which significantly affect their physiological characteristics.

If this model concerning the relationship of body image to body reactivity is generally sound, it should be applicable to other than the boundary dimension of the body image. The principal purpose of the present study was to examine the generality of the body projection model by considering a new set of variables.

The new body image dimension studied had to do with the differentiation between the left and right sides of the body. There are indications in the literature on laterality (9, 10, 13, 18), stuttering (3, 12), and body delusions associated with schizophrenia (5, 14), and organic brain pathology (4, 14) that each individual, as he learns to differentiate between the right and left sides of his body, assigns definite values and significances to each side. The kinds of values which are customarily as-

[1] This project was supported by USPH Career Research Investigator Grant M-1578.

signed have not been ascertained in any systematic objective fashion. Some have speculated (5, 21) that one side is typically perceived as strong and the other as weak. Association may also be formed between sides of the body and masculinity-femininity, goodness-badness, and the acceptable aspects of the self vs. the unacceptable aspects of the self, etc. In any case, there is some basis for regarding the right-left differentiation as a potentially significant body-image dimension. According to the body projection model, it was hypothesized that the more clearly an individual distinguishes his right and left sides in his body image, the greater is the probability that he differs in physiological reactivity on the two sides of his body.

METHOD

Measure of the Right-Left Body-Image Distinction

To test this hypothesis, it was first necessary to devise a method for measuring how clearly an individual distinguishes between his two body sides. A continuum of bigness vs. smallness was selected because it seemed to embrace roughly a number of the left-right polarities about which there has been speculation in the literature (e.g., strong vs. weak, masculine vs. feminine). The specific method that was worked out for the measurement of this right-left difference involved having each *S* make a series of comparisons of the sizes of homologous fingers of the two hands. The comparisons are made while *S* wears a set of aniseikonic lenses of 4.00% power. In a group of ten cases, there were no consistent differences in actual size between the homologous fingers of the left and right hands. Furthermore, in those instances where some differences were observed, they were of the order of one or two millimeters and could not be reliably detected without a measuring scale.

The use of aniseikonic lenses was stimulated by the work of Wittreich[2] (20), who discovered that when an *S* wears the lenses and views himself in a mirror, it creates an unstable and moderately unstructured perceptual field into which it is relatively easy for one to project unconscious feelings about his body. One of Wittreich's over-all conclusions from his data was that individuals most resist the perception of novelty or distortion in those body areas about which they have the greatest anxiety or conflict.

In the present experiment, *S* was asked to don the lenses, which were set in spectacle frames, and told that they were designed to produce various distortions in the way things looked. Questions were asked concerning the spontaneous perceptual peculiarities experienced while looking around the room. When it was apparent that *S* was beginning to experience definite distortions, a series of specific perceptual alterations were demonstrated to him. These demonstrations involved the experiencing of tilt in the walls of the room, the perception of alteration in the shape of a rectangular bookcase, and the discovery of reversal in the slope of a pointer which seemed to be pointing toward the individual even though it was actually pointing away from him at a 45° angle. *S* was then told that the experimenter was interested in the kinds of distortions the lenses would produce while he was comparing the sizes of the fingers on his two hands. He was instructed to compare in turn the five sets of fingers, beginning with the two thumbs and ending with the two little fingers. He was shown how to pair each set of fingers and to make a judgment of which was larger when viewed at arm's length. It was emphasized that the differences involved were very small and that even the most minute size discrepancy between any two fingers should be reported. The intent was to indicate that *S*'s judgment was not reality bound and that he should try to experience a difference whenever possible. It was expected that the person with a well crystallized body-image distinction between the right and left body sides would tend to perceive differences in finger sizes and that the differences would be consistently in a given direction. On the other hand, the person with poorly articulated right-left body-image distinctions was expected to be relatively conflicted and anxious about the size judgments and therefore either unable to experience differences at all or able to experience only inconsistent differences. *S*'s score for the aniseikonic judgments could vary at one extreme from seeing no differences between any of the sets of fingers to the other extreme of seeing each of the five fingers on a given body side as larger than the corresponding fingers on the opposite side.

Measure of Right-Left Body Reactivity

The GSR was selected as the index of reactivity in the two body sides. It was chosen primarily because it seemed to lend itself better than other measures to carefully controlled bilateral recordings. In order to compare the physiological reactivity of the right and left sides of the body, it was necessary to use a technique which would permit recordings that could be compared with full confidence that any differences would not be a function of differences in the apparatus associated with the measurement process on either side. The method of GSR recording was of the exosomatic type. With a constant current of 20 microamperes, resistance changes across the skin were picked up as changing potential and fed into a Grass EEG power amplifier.[3] This type of recording produces a write-out that is more a derivative of the GSR than a true wave form. But calibration against DC recording showed a linear relationship between the two. Recordings from the right and left side were made simultaneously on

[2] Wittreich, W. J. Body image and psychopathology: A pilot study. Unpublished report to Office of Naval Research, 1954. Wittreich, W. J. and Grace, M. Body image development. Unpublished progress report to Office of Naval Research, 1954.

[3] It is acknowledged with gratitude that Robert Edelberg devised and constructed the entire GSR recording system.

TABLE 1

DISTRIBUTION OF SIZE JUDGMENTS IN WHICH A MAJORITY OF FINGERS ON ONE HAND JUDGED AS LARGER THAN FINGERS ON OTHER HAND

	Right-handed			Left-handed		
	Male	Female	Total	Male	Female	Total
Majority of rights larger[a]	19	13	32 (59%)	2	6	8 (33%)
Majority of lefts larger	5	0	5 (9%)	3	5	8 (33%)
No majority	11	6	17 (31%)	6	2	8 (33%)

[a] Three or more fingers on a given hand judged as larger.

TABLE 2

DISTRIBUTION OF SUBJECTS ON SIZE JUDGMENTS CONCERNING RIGHT- AND LEFT-HAND FINGERS

	Right Fingers Larger		Left Fingers Larger	
	Right-handed	Left-handed	Right-handed	Left-handed
Number judging above median[a]	25	3	22	18
Number judging at median or below	29	21	32	6

[a] The median for the combined right and left-handed groups.

separate channels against a relatively inactive upper arm. A very large lead electrode was used on the upper arm to minimize any response in that area. Recordings were taken simultaneously first from the middle fingers; second, from the forefingers; third, from the ring fingers, and, finally, from two corresponding points on the palms. The currents to each finger were carefully matched. Also, a common calibrating resistor in series with the upper arm checked the equivalence of the sensitivities of the two circuits. The area of recording from each site was made equivalent by using pieces of tape with holes of ¼-in. diameter punched into them. The piece of tape was fixed on the recording site; a small quantity of electrode jelly was applied to the skin area enclosed by the hole; and a lead electrode was taped into place over the hole. At least two minutes of recording were obtained from each set of sites. Longer periods were monitored if it was difficult to elicit GSRs from particular *S*s. If *S* did not spontaneously produce responses, he was aroused by making loud noises, asking him to grit his teeth, suddenly touching him, etc. For each *S*, there was finally available a continuous side-by-side record of GSRs for each set of right-left sites. The amplitude of response occurring at two homologous sites at a given time could then be determined and compared. So far as can be ascertained, such simultaneous balanced recording of bilateral GSR responsivity for the purpose of studying body reactivity gradients has not been previously reported.

Subjects

There were 78 *S*s, 46 male and 32 female. Ages ranged from 22 through 38, and the median educational level was 14 years. Twenty were acquaintances of the experimenter who agreed to participate, and 58 were college students who were asked to volunteer and offered five dollars for their services.

From the very outset of the experiment, it became apparent that hand dominance was an important variable. Hand dominance was evaluated in all cases. A very approximate method was used of simply inquiring which hand the *S* used for various important activities (e.g., writing, eating, athletic skill, tasks requiring unusual strength). Fifty-four *S*s (35 male, 19 female) were right-handed and 24 (11 male, 13 female) left-handed.

RESULTS

When the judgments concerning finger size were tabulated, as shown in Table 1, it was found that 59% of right-handed *S*s considered a majority of their right-hand fingers larger than their left-hand fingers. Nine per cent considered a majority of their lefts larger, and 31% gave a majority to neither right nor left. Although the tendency was to see a majority of right-hand fingers as larger than left, 40% of *S*s either reported no consistent differences between the right- and left-hand fingers or perceived the left-hand fingers as relatively larger. The left-handed *S*s, as shown in Table 1, distributed their judgments of relative finger size equally into the categories right, left, and no difference. They did not seem to attribute obvious superiority in size to the dominant hand in the same degree as the right-handed *S*s.

A chi-square comparison of the number of *S*s in each hand dominance group who judged a majority of right-hand or left-hand fingers as larger yielded a value of 5.9, significant between the .02 and .01 levels. None of the differences among sex groups were significant.

In Table 2 is presented another analysis of the distribution of judgments concerning size of right- vs. left-hand fingers in the hand dominance groups. These data are concerned with the median number of fingers judged larger or smaller on each hand. A considerably greater proportion of right-handers than left-handers perceived an above-median number of right-hand fingers as larger than left-hand fingers. The difference, as evaluated by chi square, is significant at the .01 level. Although the left-handed *S*s did not seem to attribute clear-cut superiority in size to a majority of left-hand fingers, Table 2 indicates that they

exceeded right-handed *Ss* in the proportion of left-hand fingers described as larger. This difference, as determined by chi square, is significant at the .01 level.

In evaluating the GSR data, the criterion of directionality of response was based on a comparison of the left and right amplitudes given at each site. All corresponding responses in which a difference of at least 80 ohms occurred were noted. This level of difference was chosen because it represented 2 mm. on the measuring scale and was the smallest that could be reliably measured with a compass. Even in the maximum possible response, 2 mm. still represented 8% of the measured range. If the differences occurring for a given site were such that there were twice as many favoring one side as the other, and if at least ten differences were present, then it was considered that a directionality of response for that site had been established. Further, when *S* manifested this sort of directionality in favor of one side for three out of the four sites, he was designated as having a definite GSR gradient in the given direction. If less than three of the four sites were consistently in one direction, it was judged that a definite right-left difference was not present. For 43 of the 54 *Ss* who showed a definite directionality of response, the ratio of predominant differences to nonpredominant differences at each site was 3 to 1 or greater. Also, for 26 of these 54 *Ss*, all four of the sites showed predominant responsivity in a given direction. The reliability of the directionality of differences was explored by retesting three right-handers after intervals ranging from two weeks to two months. In all three instances, the retest directionality of response was exactly the same as that obtained for the first testing.

As can be seen in Table 3, individuals tend to show definite gradients of GSR reactivity, such that either one side responds to a relatively greater degree than the other. Only 24 of the 78 *Ss* failed to evidence a clear-cut gradient. This finding is in line with Von Helmut Baitsch's (1) report that in serial exploration of the skin resistance of the fingers on the two hands, 60 to 80% of his *Ss* had higher resistance on the right side than on the left.

There is a tendency suggested in Table 3 for right-handed women to exceed right-handed men in the relative frequency with which they manifest a definite left directionality in their GSR responses. But the chi-square value, when corrected for discontinuity, falls only at the .10–.05 level.

TABLE 3
DISTRIBUTION OF GSR RESPONSIVITY DIRECTIONALITY IN RIGHT-HANDED AND LEFT-HANDED SUBJECTS

	Right-handed			Left-handed		
	Male	Female	Total	Male	Female	Total
Right responsive	7	5	12	5	2	7
Left responsive	14	12	26	2	7	9
No clear gradient	14	2	16	4	4	8

In order to check the basic hypothesis that difference in physiological reactivity between the two body sides is a function of how definitely the two sides are differentiated in the body image, it was necessary to relate the aniseikonic finger size judgments to the GSR indices of response directionality. This was done for each of the hand dominance groups separately because it was apparent by inspection that a different pattern of relationship occurred in each. (Incidentally, the relationship between GSR directionality and lateral dominance was not in itself significant.) Within the right-handed group, the median number of right-hand fingers judged as larger than left-hand fingers was 3. The median number of left-hand fingers judged as larger was 0. Further, the median number of right-hand fingers judged larger was 2 in the group showing no distinct GSR right-left differences, 2 in the group characterized by right GSR directionality, and 4 in the group typified by left GSR directionality. *Ss* in the total right-handed group were divided into those above the median in number of right-hand fingers judged as larger and those below the median in this respect; this categorization was related by means of chi square to left directionality of GSR response vs. either no directionality or right directionality. Those above the median in their aniseikonic differentiation of their right and left sides showed a greater tendency toward left GSR directionality than did those below the median. Chi square was significant beyond the .001 level.

A word is in order concerning the manner in which GSR directionality was categorized. When a GSR gradient was present, it usually involved relatively greater reactivity on the

side judged as small than on the side described as large. Categories were established accordingly which treated left reactivity as the norm and other reactivity patterns as somehow deviant.

In the left-handed group, the median number of left-hand fingers judged larger than right-hand fingers was 2. The median number of right-hand fingers judged larger than left-hand fingers was 1. The median number of left-hand fingers judged as larger was exactly the same in the GSR right reactive, left reactive, and no directionality categories. *S*s in the left-handed group were divided into those above the median in number of left-hand fingers judged as larger and those below the median in this respect. Chi square was then utilized to determine whether these groups differed with respect to the right directional GSR category vs. the combined category of either no directionality or left directionality. The results were not significant. The GSR categories which were employed duplicated the logic of those used in the analysis of the right-handed group. When the aniseikonic judgments in the left-handed group were related to other modes of categorizing GSR reactivity, only change relationships were found.

The over-all results indicate, then, that in the right-handed group those who make the clearest aniseikonic distinction between right and left sides are those most likely to show GSR reactivity on the left side. But right-handed *S*s who distinguish poorly between their right and left sides are likely to be right reactive or to have no definite gradient at all. Among the left-handers, there was no evidence of a relationship between the aniseikonic lens judgments and directionality of GSR response.

DISCUSSION

How has the original hypothesis fared? Originally, there was no consideration of the possibility that laterality might complicate the relationships under study. It would appear, however, that only when laterality is taken into account is there evidence for the relationship between body image and body reactivity which was asserted. One may conjecture from the pattern of results that the body-image variable defined by the aniseikonic lens technique exerts its influence upon bodily reactivity within the setting of gradient effects previously established in the process of developing lateral dominance. In situations uncomplicated by body-image variables, right-hand dominance may be associated with left directionality of GSR response and left-hand dominance with the opposite GSR directionality. The experiences which determine the clarity of differentiation in the body image of the right and left body sides may occur after the basic laterality gradient has been built up, producing an impact on right vs. left reactivity within the context of the pre-existing laterality gradient. For the average person, the basic process of right-left body-image differentiation possibly involves the assigning of different values to the two body sides in such a way that attributes of bigness are projected on to the dominant side and attributes of smallness on to the other. If for some reason the division of attributes cannot be made clearly or if bigness is assigned to the nondominant rather than to the dominant side, there may be a disruption or even a reversal of the existing reactivity gradient based on laterality. At this point one can only speculate about the process involved.

The more immediate question is that of why the body-image variable is significantly related to directionality of GSR reactivity in the right-handed group but not in the left-handed group. In a sense, this question, which concerns a contrast between right and left-handed persons, is a familiar one in the literature concerned with laterality. Several studies have found right-handed persons to be more homogeneous and consistent with regard to the types of variables under consideration than left-handed persons (2, 10). An excellent and particularly pertinent study in this respect is one reported by Wishner and Shipley (19). They were interested in the effects of laterality upon the right-left directionality of autokinetic movement. They found that right-handed *S*s exhibit a consistency tendency to perceive movement toward the right, whereas left-handed *S*s evidence no consistency in either direction. It seems that left dominance is almost always far less definite than right dominance and that tension patterns associated with laterality are more irregular in the left-handed than in the right-handed. On this basis, one may speculate that the relationship between right-left body image differentiation and right-left reactivity is relatively attenuated

in the left-handed group because of the relatively diffuse and complicated character of left dominance.

Another important question has to do with the tendency for GSR reactivity to be greater on the body side perceived as the smaller. The model for explanation that comes to mind is based on the idea that the side experienced as larger (and which in right-handers with a definite GSR gradient is usually the dominant side) is assigned the function of guiding and controlling motor responses to problem situations. The executive monitoring function of one hand in an on-going task is, of course, a well confirmed phenomenon (10). The integrity and stability of this set for directive action may be of considerable psychological importance. Therefore, mechanisms may be developed for preventing surges of excitation which could disrupt or significantly alter the set. This possibility suggests the hypothesis that disruptive excitation is "drained off" or channeled into the nondominant side, where its effects are relatively less disorganizing. Thus, the nondominant side is likely to show relatively greater response to many kinds of excitation than is the dominant or "large" side. This formulation implies an optimum homeostatic[4] relationship between the two body sides. The tendency for GSRs to occur on the "small" side thus reflects the relatively greater degree of extraneous excitation diverted there in order to maintain the optimum set on the other side.

Previous work by Schoen and Scofield (15), Travis and Herren (17), and Jasper (11) lends support to this model. Schoen and Scofield found that when the eyes of a right-handed person were shifting from one fixation point to a new target, the left eye responded first, "snapping" to its new position and sometimes overshooting, as contrasted to the right eye, which proceeded more gradually and in a more coordinated fashion. Travis and Herren and Jasper observed that in situations where right-handed individuals were instructed to perform some simple task quickly and simultaneously with both arms, the left arm tended to respond first. These three studies point up the reality of a right-left difference in motoric reactivity and favor the view that the dominant side is characterized by a stable set which facilitates control but which also inhibits the spontaneity of that side as compared to the nondominant side. Jasper (11, p. 161) proposes that right-left differences in cortical organization may be responsible:

> The dominance of the cortex in the control of reflex movement is a commonly accepted principle for the explanation of the more highly integrated activity of the central nervous system. If this principle is applied to lateral dominance the tendency for the non-dominant side to lead in attempted simultaneous movement may indicate a greater cortical control ("inhibition") of the movements of the so-called dominant side which is only a counterpart of the more highly perfected coordination of movement on this side.

But the relationship found here between the right-left body image index and right-left GSR reactivity indicates that the differential in lateral response referred to by Jasper is not simply a function of lateral dominance factors. It may also be influenced by the manner in which the individual has evolved a right-left division in his body image.

The general findings essentially support the hypothesis tested. It is, of course, true that the hypothesis was oversimplified and neglected the laterality factor. However, what is outstanding is that within the right-handed group, a highly significant linkage (.001 level) was observed between the individual's body-image pattern of right-left differentiation and his pattern of right vs. left physiological reactivity. The significance of these findings increases when they are considered in the context of previously reported work (8, 9) which demonstrated that the patterning of physiological reactivity in the exterior vs. the interior of the body may be predicted from the manner in which the individual conceives of his body boundaries. At two different levels, then, body-image variables seem predictive of physiological gradient variables. The possibility is impressive that body-image concepts may prove to be one of the best means for bridging the gap between psychological and psychodynamic formulations concerning the individual and his patterns of physiological reactivity. The concepts and techniques described in the present paper lay the groundwork for investigating with considerable precision the nature of various "psychosomatic" relationships.

[4] Acknowledgement is made to Neil Burch for his helpful suggestions concerning the issue of lateral homeostasis.

SUMMARY

The hypothesis was proposed that the more definitely an individual distinguishes the right and left sides of his body in his body image, the more likely he is to manifest differential reactivity in the left vs. right sides of his body. Body-image differentiation of right and left sides was studied by means of a technique requiring *S*s to compare homologous body parts while viewing them through aniseikonic lenses. Right vs. left body reactivity was measured by simultaneous and balanced GSR recordings from the two body sides. It was found that among the right-handed *S*s those clearly distinguishing their right and left body sides were likely to show a GSR gradient such that the left side was more reactive than the right. Those without a clear body-image distinction between the right and left body sides either showed no GSR gradient or were predominantly right reactive. Among the left-handed *S*s there was no observable relationship between degree of right-left body-image differentiation and right vs. left GSR. It is considered that the data basically support the hypothesis tested.

REFERENCES

1. BAITSCH, H. Über Geschlechts- und Seitendifferenzen im "Niveau"-Elektrodermatogramm. *Conf. neurol.*, 1954, **14**, 88–100.
2. BAUER, R. W., & Wepman, J. M. Lateralization of cerebral functions. *J. Speech Dis.*, 1955, **20**, 171–177.
3. BOLIN, B. J. Left-handedness and stuttering as signs diagnostic of epileptics. *J. ment. Sci.*, 1953, **99**, 483–488.
4. CRITCHLEY, M. *The parietal lobes*. London: Edward Arnold, 1953.
5. FENICHEL, O. *The psychoanalytic theory of neurosis*. New York: Norton, 1945.
6. FISHER, S., & CLEVELAND, S. E. The role of body image in psychosomatic symptom choice. *Psychol. Monogr.*, 1955, **69**, No. 17 (Whole No. 402).
7. FISHER, S., & Cleveland, S. E. An approach to physiological reactivity in terms of a body-image schema. *Psychol. Rev.*, 1957, **64**, 26–37.
8. FISHER, S., & CLEVELAND, S. E. *Body image and personality*. Princeton, New Jersey: Van Nostrand, in press.
9. HEWES, G. W. Lateral dominance, culture, and writing systems. *Hum. Biol.*, 1949, **21**, 233–245.
10. HILDRETH, G. The development and training of hand dominance: IV. Developmental problems associated with handedness. *J. genet. Psychol.*, 1950, **76**, 39–100.
11. JASPER, H. H., & RANEY, E. T. The physiology of lateral cerebral dominance. *Psychol. Bull.*, 1937, **34**, 151–165.
12. LINDSLEY, D. B. Bilateral differences in brain potentials from the two cerebral hemispheres in relation to laterality and stuttering. *J. exp. Psychol.*, 1940, **26**, 211–225.
13. MARTIN, K. L. Handedness: A review of the literature on the history, development, and research of laterality preference. *J. educ. Psychol.*, 1952, **45**, 527–533.
14. SCHILDER, P. *The image and appearance of the human body*. New York: International Univer. Press, 1950.
15. SCHOEN, Z. J., & Scofield, C. F. A study of the relative neuromuscular efficiency of the dominant and non-dominant eye in binocular vision. *J. gen. Psychol.*, 1935, **12**, 56–181.
16. THASS-THIENEMANN, T. Left-handed writing. *Psychoanal. Rev.*, 1955, **42**, 239–61.
17. TRAVIS, L. E., & HERREN, R. Y. Studies in stuttering. V. A study of simultaneous antitropic movements of the hands of stutterers. *Arch. Neurol. Psychiat.*, 1929, **22**, 487–494.
18. WILE, I. S. The relation of left-handedness to behavior disorders. *Amer. J. Orthopsychiat.*, 1932, **2**, 44–57.
19. WISHNER, J., & SHIPLEY, T., JR. Direction of autokinetic movement as a test of the "Sensory-Tonic-Field" Theory of perception. *J. Pers.*, 1954, **23**, 99–106.
20. WITTREICH, W. J., & RADCLIFFE, K. B., JR. The influence of simulated mutilation upon the perception of the human figure. *J. abnorm. soc. Psychol.*, 1955, **51**, 493–495.
21. WOLFF, W. *The expression of personality*. New York: Harper, 1943.

Received August 3, 1957.

TIME ORIENTATION AND INTERPERSONAL RELATIONS OF EMOTIONALLY DISTURBED AND NORMAL CHILDREN[1]

ANTHONY DAVIDS AND ANITA NEGRIN PARENTI

Brown University and Emma Pendleton Bradley Hospital

THE aim of this research was to investigate time orientation and interpersonal relations in a group of hospitalized emotionally disturbed children and in two groups of normal children. Time orientation in imaginal processes was studied by a projective method, and interpersonal relations were defined by measures of stability of friendship patterns. These two significant areas were previously investigated independently (1, 5, 7). The present study was designed to replicate and extend these previous studies and to relate and unify them.

In a study of time orientation and social class, LeShan (6) found that lower-class children tend to be present-oriented, while middle-class children tend to be future-oriented. In a related study, Barndt and Johnson (1) compared time orientation in groups of delinquent and nondelinquent boys who were equated for socioeconomic status. They found the delinquents to be significantly more present-oriented than were the nondelinquents. In both of these investigations, it was speculated that emotionally disturbed individuals have short-time perspectives. On the basis of a recently reported experiment with psychotic adults, Wallace stated that "... future time perspective is influenced by the schizophrenic process to such an extent that both the length of the future time span and the organization of its contents are significantly reduced for a sample of schizophrenic patients as compared to a group of normal controls. Additional research is needed, however, before such a finding can be considered to have been conclusively demonstrated" (8, p. 245).

Two studies on stability of friendship patterns, carried out by Thompson and Horrocks (5, 7), are also directly relevant. While other inquiries focused on time orientation as assessed by responses to a projective technique, Thompson and Horrocks investigated fluctuations in friendship choices in everyday social relations. With both rural and urban children, they found that in 11-year-olds there was considerable fluctuation in friendship choices over a two-week period of time. However, with increasing age there was an increasing degree of stability in friendship patterns. Thus, it would appear that stability of friendship patterns is a correlate of emotional stability.

Goldfarb (4) showed that children who are emotionally deprived in their early years and who are reared in institutional settings tend to have deficient time orientation. He argued that this deficiency is definitely a factor in their limited foresight and their inability to envision the consequences of their acts and is even a major cause of their social maladjustment.

It appears, then, that time orientation is crucial both to emotional and social adjustment. The present study examines this notion empirically.

METHOD

Subjects

The experimental *Ss* in this study consisted of a group of 30 hospitalized emotionally disturbed boys undergoing residential treatment at the Emma Pendleton Bradley Hospital in Riverside, Rhode Island. They ranged in age from 7 years, 5 months, to 13 years, 4 months, with a mean age of 10 years, 10 months. The children in this group were suffering from severe behavior disorders, neuroses, or psychoses. Most of them had been at the Bradley Hospital for a period of not less than eight months, and the average length of hospitalization was approximately two years. The boys represented a heterogeneous sample with respect to socioeconomic status, but most of them were from middle- and lower-class families.

[1] We wish to express our appreciation to the director and the child care staff of the Bradley Hospital for their cooperation in our work with the disturbed children. We are also indebted to Edward E. Martin, superintendent, and Louise Favorite, psychologist, for their cooperation in our study of children in the East Providence, Rhode Island public schools. And Murray Ortof, director, was very helpful in our study of children at Camp Hurley in Kingston, New York. We are also grateful to Sylvia Topal, staff social worker at the camp, and Johanna Walter, psychiatrist at Bradley, who provided clinical ratings of the children. This research was facilitated by a summer undergraduate research stipend awarded to Anita Negrin Parenti from the Social Science Research Council, and by Research Grants M-1722 and M-2087 awarded to Anthony Davids from the National Institute of Mental Health, Public Health Service.

One group of 33 control *S*s was drawn from a population of children attending a summer camp for a period of three weeks. The camp (Camp Hurley, in Kingston, New York) took in children from all socioeconomic and ethnic groups. Although social workers were employed as supervisors to the counselors, who lived and worked in direct contact with the children, the camp was for normal youngsters. The boys studied ranged from 7 years, 11 months, to 13 years of age, with a mean of 10 years, 10 months.

Another control group of 22 boys was obtained from the Grove Avenue School in East Providence, Rhode Island. The children were in the fourth, fifth, and sixth grades. They ranged from 9 years, 8 months, to 12 years, 4 months of age, with a mean of 10 years, 10 months.

Thus the *S*s were comparable for age, and although they were not, strictly speaking, equated for socioeconomic status, there is no reason to believe that any systematic class differences existed among the groups.

Time Orientation in Fantasy

The story-completion technique. The procedure followed was previously employed and described by Barndt and Johnson (1) and by Wallace (8). The *S* was given the following instructions: "I want to see what kind of a story you can tell. I'll start a story and then let you finish it any way you want to. You can make it any kind of a story you wish. Do you understand? I'll start it now." All four of the following stories were administered to the Bradley sample, the first two to the camp sample, and only the first one to the school sample. The first story began as follows: "About three o'clock one bright sunny afternoon in May, two boys were walking along a street near the edge of town. . . ." The second story began: "Ten o'clock one morning Al met his friend Jerry near the center of town. . . ." The beginning of the third story was: "Two boys were walking along the street three o'clock one sunny afternoon. One boy said to the other. . . ." And the fourth story began: "Johnny and Jimmy met each other ten o'clock one morning. . . ." After reading the beginning of each story to the child, *E* told him, "Now you start there and finish the story any way you want to." *E* then recorded the child's story verbatim. After finishing the story, the child was asked, "What time was this?" or "How long did the story take to happen?" In the cases where children were unable to report the length of time covered by their stories, *E* estimated the time span, but this occurred very infrequently.

In the Bradley sample, the first and second stories were administered during the first interview and the remaining two after two weeks. In the camp sample, only the first story was given during the first interview; the second was given three weeks later.

Scoring of stories. The scoring system described by Barndt and Johnson (1) was used. The scoring categories were as follows:

1. Under one hour
2. One hour to under five hours
3. Five hours to under twelve hours
4. Twelve hours to under one week
5. One week to under three months
6. Three months or more

The amount of time covered in the stories could be scored very accurately, and agreement between raters was unity.

Reliability of measures. Using the six-point scale described above, a product-moment correlation of .75 between the first and second stories was obtained for the Bradley sample. Correlations of .66 and .73 were found between the second and third and the third and fourth stories. In the camp sample, a correlation of .93 was obtained between the first and second stories. These coefficients indicate adequate consistency in these measures of time orientation. Only scores from the first story were used in the analyses reported here.

Stability of Interpersonal Relations

The friendship fluctuation method. In assessing friendship patterns, the method used by Thompson and Horrocks (5, 7) was adapted. The child was told: "Please tell me the names of your three best friends, or the children you like best at Bradley (camp or school). Tell me the one you like best, then the second best, and then the third best." The children were asked to name their friends on four different occasions, with a one-week interval between each interview. At each of the subsequent interviews, the original instructions were repeated and the following instructions were added: "The people you tell me may or may not be the ones you liked best the last time. Be sure to tell me the three you like now."

Scoring of friendships. To score the friendship patterns, a point system was taken from Thompson and Horrocks (5, 7) with a few modifications. One point was given for each change in position of a friend or friends. For example, if John were selected during the first interview as the very best friend of *S* and if during the next interview he was selected as *S*'s third best friend, two points were awarded. If John were selected as *S*'s second best friend during the second interview, one point was awarded. Seven points were awarded for the selection of a best friend who did not appear in the choice pattern during the previous interview. Similarly, six points were awarded for the selection of a second best friend and five for the selection of a third best friend who did not appear in the previous choice pattern.

The highest score that could be obtained by retaining the same friends but switching the order of preference was four. It was felt that a higher degree of instability was indicated by choosing an entirely new friend or friends than was indicated by switching the order of preference of previously named friends. More points were therefore given for the appearance of a new friend than for the maximum changing of preferences.

By means of this scoring system, it was possible to obtain several friendship fluctuation scores for each child: (*a*) over three weekly periods, (1–2, 2–3, and 3–4); (*b*) over two two-week periods, (1–3, and 2–4); and (*c*) over one three-week period, (1–4). Except where otherwise noted, however, each child's friendship fluctuation score represents the mean of the three weekly periods. This score was used in computing all correlations. For purposes of comparison, the friendships were rescored according to the same method used by Thompson and Horrocks (5, 7). This procedure entailed

scoring the change in friendships over a two-week period. Four points were awarded for the appearance of a new best friend in the choice pattern, three points for the appearance of a new second best friend, and two points for the appearance of a new third best friend. In addition, one point was awarded for each change in position of a friend or friends. In the statistical treatment of the data, it was felt that it was better to consider the friendship fluctuations of weeks two-to-four in the camp group as being comparable to weeks one-to-three in the Bradley and school groups. This was done because the children in the camp group had not known each other for more than one day when the first friendship measure was taken, while those in the Bradley and school groups had known one another for some time prior to the study. However, this method was used only in comparing the camp group to the children studied by Thompson and Horrocks (5, 7).

Measurement of Emotional Disturbance

The rating scales. Rating scales were constructed for each of the following eight personality dispositions: sociocentricity, optimism, trust, egocentricity, pessimism, distrust, anxiety, and resentment. Of these traits, the first three are positive or socially desirable; the last five are negative or socially undesirable (2). In addition to ratings on the individual traits, a scale was constructed on which *S* could be rated in terms of the degree of his "general emotional maladjustment." Detailed definitions of the variables were included on the mimeographed sheets containing the rating scales. The raters were instructed to judge each *S* in comparison with other children of his age and sex in his social group. The ratings were made on six-point scales running from "much less than most" to "much more than most," with a score of one being assigned to the lower extreme and a score of six assigned to the higher extreme. Thus, it was possible to utilize scores on the individual dispositions, sums on the three positive or on the five negative dispositions, or scores on the general rating.

Reliability of ratings. Ratings were obtained for the Bradley sample and the camp sample. The children at Bradley were rated by a staff clinical psychologist and a staff psychiatrist. The children at camp were rated by counselors and a professional social worker. The product-moment correlation between the ratings made by the psychologist and psychiatrist on general emotional maladjustment for the Bradley sample was .77, while the comparable correlation between the counselors and social worker at camp was .54. In the statistical comparisons contained in this report, the ratings by the psychologist (Bradley) and by the social worker (camp) were employed.

HYPOTHESES AND RESULTS

The hypotheses to be tested in this experiment were derived from the theoretical discussions and reports of previous investigators who have found present orientation and instability of friendships to be correlates of immaturity and emotional maladjustment.

Hypothesis I: *Disturbed children are more present-oriented in fantasy than are normal children.* It is evident from Table 1 that this hypothesis was not confirmed. Actually, the disturbed group received a slightly higher, though nonsignificant, score on future orientation than did the two normal groups.

Hypothesis II: *Disturbed children show less stable friendship patterns than do normal children.* The findings in Table 1 indicate that the disturbed sample was characterized by a significantly higher friendship fluctuation score than either of the control groups, while the normal groups did not differ from one another. Thus, it appears that the patterns of interpersonal relations in the disturbed group are significantly more unstable than in the normal groups. The percentage of change in choice of best friend over a two-week period was also

TABLE 1

COMPARISON OF GROUPS ON STORY-COMPLETION AND FRIENDSHIP FLUCTUATION SCORES

Measure	Bradley ($N = 30$)		Camp ($N = 33$)		School ($N = 22$)		Bradley vs. Camp		Bradley vs. School		Camp vs. School	
	Mean	S^2	Mean	S^2	Mean	S^2	F	t	F	t	F	t
Friendships (Mean)	8.13	15.62	5.39	8.57	6.04	16.52	1.82	3.22**	1.06	1.86*	1.93*	.65
Friendships (1–2)	8.97	22.48	6.18	13.09	6.23	22.56	1.79	2.86**	1.04	2.04*	1.72	.04
Friendships (2–3)	8.00	20.90	5.48	17.45	6.91	33.13	1.20	2.36*	1.59	.76	1.90	1.07
Friendships (3–4)	7.43	29.22	4.52	12.57	5.00	18.57	2.32*	2.51**	1.57	1.74*	1.48	.45
Friendships (1–3)	9.07	21.93	7.00	18.63	7.45	24.45	1.18	1.88*	1.11	1.21	1.31	.36
Friendships (2–4)	9.80	31.48	5.79	20.48	6.73	27.06	1.54	3.23**	1.16	2.02*	1.32	.71
Friendships (1–4)	9.60	21.90	7.61	18.93	6.23	16.47	1.16	1.75*	1.33	2.70**	1.15	1.18
Friendships (1–3) Thompson Scoring	4.73	5.44	3.73	4.83	3.95	6.52	1.13	1.79*	1.20	1.15	1.35	.33
Story scores	2.93	2.69	2.39	1.31	2.86	1.36	2.05	1.50	1.98	.17	1.04	1.49

* Significant at the .05 level with a one-tailed test.

** Significant at the .01 level with a one-tailed test.

computed. The percentages obtained were 47, 42, and 59 for the Bradley, camp, and school samples respectively; there are no significant differences among these values.

Hypothesis III: *There is an association between time orientation in fantasy and stability in friendship patterns.* In the Bradley group, there is a significant positive correlation of .61 between friendship and story scores, while in the two normal groups there is no relation between the two measures. This indicates that with the emotionally disturbed *S*s, the more present-oriented the child is on the story-completion method, the more stable his friendship patterns tend to be.

Hypothesis IV: *Degree of general emotional maladjustment is correlated with time orientation and with stability of friendship patterns.* In the Bradley sample, there is no relation between these variables. In the camp sample, there is only a nonsignificant positive correlation of .22 between the ratings on general emotional maladjustment and story scores. In the normal sample, there is a significant positive correlation of .33 between the ratings on maladjustment and friendship fluctuation scores. This indicates that the more disturbed the child is, within the normal sample of children, the more likely he is to change his friends from week to week.

Hypothesis V: *Ratings on the positive personality traits are correlated with future orientation in fantasy and stability of friendship patterns, while ratings on the negative personality traits are correlated with present orientation and instability of friendships.* The ratings on the three positive traits (sociocentricity, optimism, trust) were summed to arrive at a general rating of good adjustment, as opposed to the previous rating on general emotional maladjustment used in testing Hypothesis IV. As can be seen from Table 2, Hypothesis V was not confirmed. For the Bradley sample, this positive sum bore no relation to either the story or friendship scores. In the normal sample, however, there is a significant correlation of −.41 between the sum of the positive traits and story scores, indicating that the more present-oriented a child is, the more likely he is to be rated high on positive traits. The correlation of −.22 between the sum of the positive traits and the friendship scores is nonsignificant.

Correlations between ratings on the individual personality traits and the story and friendship scores are shown in Table 2. The results indicate no significant relations between the story scores and the individual traits in the Bradley sample. In the camp sample, there is a significant correlation of .50 between pessimism and story scores, indicating that the children who were judged to be pessimistic tended to be future-oriented on the fantasy measure. The rating on resentment also correlated significantly (.41) with story scores, suggesting that the more resentful children were more future oriented And the significant correlation of −.43 between trust and story scores implies that the more trusting children tended to be more present-oriented in fantasy.

Of the correlations between personality traits and friendship fluctuations, the only significant coefficient (.48) in the Bradley sample suggests that children judged clinically to be high on resentment evidence greater fluctuations and instability in their choice of friends from week to week. In the camp sample, the significant correlation of −.30 is consistent with the notion that the more sociocentric a child is, the less fluctuating are his friendship patterns.

While these findings lack consistency, they

TABLE 2

CORRELATIONS BETWEEN MEASURES OF TIME ORIENTATION AND RATINGS ON PERSONALITY TRAITS

Group and Measure	Egocentricity	Pessimism	Distrust	Anxiety	Resentment	Sociocentricity	Optimism	Trust	Sum of Positive Ratings
Bradley (N = 30)									
Story scores	.02	−.19	.03	.10	.15	−.08	−.06	−.03	−.08
Friendship scores	−.06	−.10	.00	.03	.48**	−.07	−.11	−.08	−.10
Camp (N = 33)									
Story scores	.28	.50**	.33	.15	.41*	−.31	−.29	−.43*	−.41*
Friendship scores	.24	.22	.27	−.10	.25	−.30*	.02	−.25	−.22

* Significant at the .05 level.

** Significant at the .01 level.

are characterized by certain noteworthy trends. The ratings on the negative traits appear to be positively correlated with both story and friendship scores, while the ratings on the positive traits are, with only one exception, negatively correlated with these variables. This indicates that the more well-adjusted children tend to show less change in their friendships and tend to tell stories that cover a shorter time span.

In addition to tests of the hypotheses, some direct comparisons were made between the results obtained here and those reported by previous investigators. A comparison between the results of this study and Barndt and Johnson's (1) in their study of 17-year-old delinquent and nondelinquent boys was possible because the story beginnings and the scoring system were identical. The results presented in Table 3 show that all of our three groups of 11-year-olds were more present-oriented than their group of 17-year-old nondelinquents, whereas there were no differences between our groups and their delinquents.

The mean friendship fluctuation scores of the two normal groups in the present study, when computed for a two-week period by the exact method used by Thompson and Horrocks (7) in their study of urban children, did not differ significantly from the scores reported by these investigators. The degree of friendship fluctuation for the disturbed group, however, was greater than that reported for their normal 11-year-old boys. These investigators found approximately a 50% change, at this age level, in choice of best friends over a two-week period. As noted above, the change within each of the three groups in the present study was also approximately 50%.

TABLE 3

COMPARISON OF PRESENT THREE GROUPS WITH GROUPS STUDIED BY OTHER INVESTIGATORS

Group	M	S^2	Bradley		Camp		School	
			F	*t*	*F*	*t*	*F*	*t*
Barndt & Johnson Delinquents ($N = 26$)	2.8	1.96	1.37	.32	1.50	1.24	1.44	.16
Barndt & Johnson Controls ($N = 26$)	3.6	1.69	1.59	1.68*	1.29	3.78**	1.24	2.05*
Thompson & Horrocks 11-year-olds ($N = 21$)	3.48	4.62	1.18	2.03**	1.13	.38	1.41	.70

* Significant at .05 level.
** Significant at .01 level.

DISCUSSION AND CONCLUSIONS

While there were no significant differences between normal and disturbed boys with respect to time orientation as measured by the story-completion technique, a tendency toward future orientation in the camp group was found to be positively associated with the personality traits of pessimism and resentment. Future orientation was also found to be negatively associated with the sum of the ratings on the three positive personality traits. These findings suggest the need for a re-examination of the hypothesis that future orientation in the fantasies of 11-year-old children is a sign of a mature and healthy adjustment pattern. It seems, to the contrary, that the better adjusted children tended to be more present-oriented than those judged as less well-adjusted emotionally.

Further evidence in support of this revised hypothesis can be derived from the relation between friendship patterns and story scores in the disturbed sample. Thompson and Horrocks (5, 7) demonstrated developmentally with normal children that there is increasing stability of friendships with increasing chronological age. Thus, low friendship fluctuation scores may be considered an index of maturity. In the Bradley sample, a considerable degree of association was found between the friendship fluctuation scores and the story-completion scores, indicating that the children with the more stable friendships tended to be more present-oriented in fantasy.

In view of the evidence at hand, some tentative conclusions are appropriate with regard to time orientation in imaginal processes. A high degree of future orientation in 11-year-old children appears to be associated with inadequate personal and social adjustment. It seems plausible that a high degree of future orientation in fantasy indicates that the child is dissatisfied in the present and is investing a good deal of his energy in fantasy about the future. In other words, children who find their everyday lives unsatisfying and who try to escape their present situation by means of fantasies about the future may well be more disturbed, while the more stable children invest more of their energy and thoughts in

present life situations. In considering these possibilities, it should be emphasized that what may be a sign of good adjustment at one age level may well be abnormal at another. For example, the normal 17-year-olds in Barndt and Johnson's study were more future-oriented than any of our groups of 11-year-old boys, whereas the 17-year-old delinquent boys were similar to our 11-year-olds in their present orientation. Understandably, it may not be until early adolescence that a person is normally concerned more with the future than with the present. Thus, although present-oriented at 11 years or younger may signify healthy adjustment, future orientation at some later age may reflect maturity and healthy emotional adjustment. These notions are largely speculations, but further research can certainly provide more definite answers.

The findings in regard to friendship patterns were unequivocal, providing quite convincing evidence that emotionally disturbed children tend to vary their friends over short intervals of time to a greater degree than do normal children. These significant differences in friendship fluctuations were found when the present group of disturbed children was compared both with the present normal groups and with the normal children studied previously by Thompson and Horrocks. Therefore, in addition to confirming an experimental hypothesis, the finding of striking similarities between the present normal *S*s and those studied in an independent investigation serves as a replication. As we (3) and others have argued elsewhere, there is great need for the worthwhile contributions of exact replications in psychological experiments. When the replications yield similar results, then the findings obtained in modifications and extensions of the original studies become all the more meaningful. In this regard, the previous work with normals has been extended to the social relations of disturbed children.

In the camp sample, there was a significant correlation between the ratings on emotional disturbance and friendship fluctuations scores, with the more disturbed children showing a greater change in friends. While there was no association between the clinical ratings and the friendship fluctuations in the Bradley sample, this may be attributable to the fact that these boys formed a very homogeneous group in terms of emotional disturbance. The fact that they were all institutionalized because of this difficulty indicates that they all were far out on the continuum of maladjustment, and this homogeneity would be conducive to low correlations between relative standing on this variable and scores on the measure of friendship fluctuations.

At any rate, the objective evidence is in keeping with clinical impressions of unstable social relations in people who suffer from severe psychopathology. In many cases of childhood emotional disorders, the parent-child relation is unstable and unsatisfying, and it is understandable that this unhappy interpersonal pattern may well generalize into a general distrust and other negative traits that make it difficult for the youngster to establish effective relations with either peers or adults. Thus, a vicious cycle is set up, with emotional disturbance leading to disturbed social relations, and the lack of social adjustment making it impossible for the disturbed child to obtain the kinds of gratification (e.g., social approval, lasting friendships, love) that are necessary for healthy personality development. A fundamental aspect of the psychotherapeutic process with children is the establishment of a trusting, secure, warm relation with a significant adult in the hope that on this foundation a good personal and social adjustment may be built. From both a therapeutic and a scientific point of view, however, many lines of research remain to be investigated before any adequate understanding is available of the interpersonal relations that characterize emotionally disturbed children.

SUMMARY

Three groups of 11-year-old boys were studied in the following settings: (*a*) residential psychiatric treatment center, (*b*) summer camp, and (*c*) public school. A story-completion technique was used to investigate time orientation as revealed in imaginal processes. Assessment of fluctuations in friendship patterns over a three-week period was the method used to study interpersonal relations. Ratings on general emotional adjustment and several positive and negative personality traits were obtained from qualified staff personnel.

In comparison with normals, as predicted, emotionally disturbed children showed more

unstable friendship patterns. However, contrary to prediction, their story completions were not more present-oriented. In the normal camp group there was a significant positive association between ratings on emotional maladjustment and degree of friendship fluctuation. In the disturbed group, the more present-oriented a child was on the fantasy measure, the more stable were his friendships. With normal children there was no association between time orientation and the measure of interpersonal relations.

The present *Ss*' scores were compared with those obtained by Barndt and Johnson, using the identical story-completion technique, in a study of 17-year-old delinquents and nondelinquents. All of our groups were significantly more present-oriented than were their nondelinquents, while there were no differences between our groups and their delinquents.

No differences were found between friendship fluctuation scores of the present normal groups and scores obtained previously by Thompson and Horrocks in studies of normal urban children. In each of our groups, there was approximately 50% change in choice of best friend over a two-week period, a result which replicates theirs.

It was concluded that disturbed and normal 11-year-old boys do not differ in fantasy measures of time orientation, nor do they differ from older delinquents, but they are more present-oriented than older normal boys. Also, unstable social relations are characteristic of emotionally maladjusted children.

REFERENCES

1. BARNDT, R. J., & JOHNSON, D. M. Time orientation in delinquents. *J. abnorm. soc. Psychol.*, 1955, **51**, 343–345.
2. DAVIDS, A. Alienation, social apperception, and ego structure. *J. consult. Psychol.*, 1955, **19**, 21–27.
3. DAVIDS, A. The influence of ego-involvement on relations between authoritarianism and intolerance of ambiguity. *J. consult. Psychol.*, 1956, **20**, 179–184.
4. GOLDFARB, W. Psychological privation in infancy and subsequent adjustment. *Amer. J. Orthopsychiat.*, 1945, **15**, 247–255.
5. HORROCKS, J. E., & THOMPSON, G. G. A study of friendship fluctuations of rural boys and girls. *J. genet. Psychol.*, 1946, **69**, 189–198.
6. LESHAN, L. L. Time orientation and social class. *J. abnorm. soc. Psychol.*, 1952, **47**, 589–592.
7. THOMPSON, G. G., & HORROCKS, J. E. A study of friendship fluctuations of urban boys and girls. *J. genet. Psychol.*, 1947, **70**, 53–63.
8. WALLACE, M. Future time perspective in schizophrenia. *J. abnorm. soc. Psychol.*, 1956, **52**, 240–245.

Received August 7, 1957.

THE INTERACTION OF ACHIEVEMENT MOTIVATION AND ABILITY IN PROBLEM-SOLVING SUCCESS[1]

ELIZABETH G. FRENCH

Air Force Personnel and Training Research Center

THE main purpose of the present study was to test an hypothesis concerning the relationship between ability level, achievement motivation, and problem-solving performance. The experiment was also designed so that a previously established relationship (3) between achievement motivation and performance could be tested under somewhat different conditions, and the nature of that relationship could be examined further.

The present hypothesis, an outgrowth of earlier studies (1, 3), states that performance in a problem-solving situation will be related to ability level among subjects (*S*s) with high achievement motivation, but there will be little or no relationship among *S*s with low achievement motivation. The reasoning behind this prediction is that *S*s with high achievement motivation presumably will attempt to use their ability as effectively as possible, thus differences in that ability will be reflected in performance. *S*s with low motivation, on the other hand, will tend not to use their potential, so those with high ability will do little or no better than those with low ability.

The additional relationships studied concern the effect of achievement motivation on performance. A previous study (3) demonstrated that motivation was related to successful performance in a problem-solving situation where the length of time worked was determined by *S*. Although it could be shown that this relationship remained when length of time worked was partialled out, the present experiment was designed to test the hypothesis that the motivation-performance relationship exists when time worked is the same for all *S*s.

In the earlier study, also, it appeared that the relationship between motivation and problem success was not uniform over the entire motivation range. There was a steady increase in percentage of successes with increase in motivation interval up to about the median of the motivation distribution. Beyond that point the percentage of successes was roughly the same for all motivation scores. That is, any motivation level above the median was sufficient to give the maximum effect of that variable. This finding was not predicted nor expected on any theoretical grounds. The explanation could be that this is the general nature of the relation between motivation and performance or it could lie in the specific problem—its difficulty level, intrinsic interest value, etc. This experiment included in the design the possibility of testing for the existence of the described relationship under conditions slightly different from those where it was previously noted.

The specific hypotheses to be tested are:

I. Intelligence will be more highly related to success in problem-solving among *S*s with achievement motivation scores above the median than among *S*s with motivation scores below the median.

And the additional hypotheses:

II. *S*s with achievement motivation scores above the median will solve more problems in a given time period than will *S*s with motivation scores below the median.

III. Differences in motivation scores below the median of the distribution will be more highly related to problem-solving performance than will differences in motivation score above the median.

METHOD

Measure of Motivation

The measure of motivation was a projective device called the Test of Insight (2) in which the *S* is asked to provide "explanations" for briefly described behaviors presented as typical of hypothetical individuals. The number of achievement-relevant explanations make up the score.

The Problem Apparatus

The problem-solving apparatus consisted of a panel containing a horizontal row of seven lights and

[1] This report is based on work done under ARDC Project No. 7719, Task No. 17011, in support of the research and development program of the Air Force Personnel and Training Research Center, Lackland Air Force Base, Texas. Permission is granted for reproduction, translation, publication, use, and disposal in whole and in part by or for the United States Government.

another of seven switches. The wiring of the board was such that some switches operated one light; some, two; and some, three. Similarly, some lights were controlled by one switch only and some by two or three. Operation of a switch always affected all the lights to which it was connected. Thus, operation of Switch 1, controlling Lights 1 and 3 would put those lights on; subsequent operation of Switch 2, controlling Lights 1 and 2 would extinguish Light 1 (already on) and put on Light 2. The *S*s were required to create specified patterns of lights.

Subjects

The *S*s were 96 airmen undergoing basic training at Lackland Air Force Base. They were selected to cover the Armed Forces Qualifications Test (AFQT) range of 70–99, with equal numbers in the 70–79, 80–89, 90–99 ranges.

The *S*s were also selected so that equal numbers of each quartile of the achievement motivation fell in each AFQT decile. This made a 3 × 4 design with three levels of AFQT score and four levels of achievement motivation with eight *S*s per cell.

Procedure

Each day's *S*s were met as a group by *E* and given an orientation to the problems which established them as measures of general intelligence and all-round capability which gave an accurate estimation of what a man could accomplish. Immediately following this, they were given the Test of Insight, described as unrelated to the problems. Upon completion of this test, each *S* went individually to the problem-solving session.

First, *E* demonstrated the apparatus to *S* and stressed the important points, i.e., that switches could operate more than one light, and vice versa, and that the operation of a switch affected all the lights it controlled. Then *S* was permitted to operate the switches, one at a time, in order to become familiar with the switch-light combinations. At the end of about a two-minute practice period he was asked to recite these combinations to *E*. When *S* had demonstrated that he knew the combinations, *E* read off four demonstration problems together with their solutions, and *S* operated the designated switches to produce the light pattern stipulated. Then *E* read the following instructions:

> The problems you will solve will be like these you have just done. In working them out you may use the switches as often as you like but you may have only one on at a time, as you did when you were learning the combinations. In other words, you must work the solution out in your head, not on the board. When you think you have the answer, when you think you know which switches will give you the pattern of lights you want, tell me what you think it is so I can record it, and then try it out on the board the way you did the demonstration problems. If it is correct, if you get the correct lights, turn them off and go directly to the next problem. If it is wrong, turn the lights off and keep working on the same problem. Take each problem in turn, solving one before you go to the next. You will have 25 minutes to solve as many problems as you can
>
> Is it clear? You work out the solution, using the switches one at a time if you like, call out all solutions before you try them, try the solution, if it works go to the next problem, if it doesn't keep working on the same one.
>
> The problems are written on these cards, one problem to a card. They tell you the number of switches you need to use and the lights you are to get on.

The problems were written on individual 3 × 5 cards and placed on the table by the apparatus. (An example: What *three* switches will give you Lights 1, 3, and 6?) The first four were warm-up problems designed to be solved in about one minute. Their purpose was to provide more familiarity with the apparatus and to minimize practice effect in the series proper so that mean solution times would be approximately the same for all problems. The regular series consisted of 16 equal-difficulty problems. It was anticipated that no *S* could complete the series in the time allotted. Scoring and timing started with the regular series, and *S* was given a two-minute warning of the time limit. The score was the number of problems solved in the 25-minute time limit.

RESULTS AND DISCUSSION

Before the main data analyses were carried out, the time spent on the preliminary problems with respect to its relation to number of regular problems solved and the individual problem mean times were examined. The mean time required to solve the preliminary problems was 5.69 minutes. This time varied widely within cells and also varied as a function of motivation score ($F = 3.67$, $P = .05$). Since the time spent on the preliminary problems varied in the same way as time per problem in the regular series (r for preliminary problem time vs. number of regular problems solved $= -.23$), it cannot be argued that the success on the regular series was due to a longer practice time in the preliminary period. In the regular series, the mean times for the first eight problems (not enough *S*s completed more than eight to permit stable means for the remaining problems) were, in order: 2.92, 3.21, 2.63, 3.54, 2.85, 4.65, 3.57, 2.91. They are all fairly homogeneous with the exception of that for Problem 6, which apparently was more difficult than the others. The effect of this greater difficulty of that one problem will be discussed later.

The mean AFQT and achievement motivation scores were also checked for individual cells. No differences were found in motivation means across a given AFQT level or vice versa.

The problem-solving scores ranged from 1 to

12, with a mean of 5.67 and a standard deviation of 2.33. The means for the 12 cells are presented in Table 1. The four quartiles of the achievement motivation distribution are delimited by scores 0–4, 5–7, 8–10, and 11–20.

Since the first two hypotheses dealt with differences expected as a result of motivation scores above or below the median, the two top and the two bottom quartiles were combined for the analyses pertaining to these hypotheses. The summary of the analysis of variance is presented in Table 2. The *F* for the interaction between AFQT and achievement motivation score (the test of the main hypothesis) is significant at the .01 level. Tukey's gap test gave a difference of 1.04 between means as a significant gap. By this test, the means for the three AFQT levels in the high motivation group are significantly separated from each other; those in the low motivation group are not. The relationships between AFQT and performance for the two motivation groups are also described by correlations of .60 and .08 for high and low motivation respectively. The hypothesis that ability level will be related to problem-solving performance among *S*s with high achievement motivation and unrelated among *S*s whose achievement motivation was low is supported by the data.

The second hypothesis, that motivation is related to problem-solving success when time worked is constant was also clearly supported by the analysis. The *F* for achievement motivation is significant beyond the .001 level. The correlation between the variables is .58.

The support for the first two hypotheses is fairly unequivocal and needs little comment. Hypothesis III requires a little more discussion. Examination of the means in Table 1 shows that the data fail to give it support. An analysis of variance of the 12-cell table of means breaking down the variance due to motivation into that due to position with respect to the median, that due to quartile within halves of the distribution, and the interaction of the two, yields *F* ratios considerably less than one for both the quartile position and the interaction between quartile position and position with respect to median. Thus it is apparently not true that in general the relation between achievement motivation and problem-solving is closer below the median of the motivation distribution than above it. In this case, the median of the distribution provided a sharp dividing line for problem-solving scores. Few *S*s with motivation scores below the median did well and few with scores above did poorly. There was very little overlap between the distributions. But the actual score above or below the median made little difference. That is, those with very high scores did not do noticeably better than those with moderately high scores, nor did those with very low scores do more poorly than those whose scores were moderately low.

In an attempt to find a tentative explanation for the failure to demonstrate the kind of relationship that had appeared in the previous experiment the distributions of problem-solving scores were examined for the several achievement motivation groups. This examination revealed first that the difference between the combined high motivation and the combined low motivation groups was manifested immediately. All the high *S*s were successful through Problem 4 and only two failed Problem 5. It wasn't until Problem 6 was reached that an appreciable number of those *S*s—six—failed. The failures in the low groups began with the second problem. Coming now to the two low quartiles, it is clear that they do not differ through Problem 4. Seventeen third and 16 fourth quartile *S*s reached and passed that problem. But with Problem 5, the beginning of a difference is sug-

TABLE 1
MEAN PROBLEM-SOLVING SCORES BY ACHIEVEMENT MOTIVATION AND AFQT LEVEL

AFQT	Motivation Score			
	11+	8–10	5–7	4–
90–99	8.75	9.00	4.12	4.00
80–89	7.25	6.87	4.25	4.12
70–79	6.00	6.00	4.00	3.62

TABLE 2
SUMMARY OF ANALYSIS OF VARIANCE OF PROBLEM-SOLVING SCORES WITH TWO LEVELS OF MOTIVATION AND THREE LEVELS OF AFQT

Source	Sum of Squares	*df*	Mean Square	*F*	*P*
Motivation	260.04	1	260.04	120.39	<.001
AFQT	39.14	2	19.57	9.06	<.01
M × AFQT	29.65	2	14.83	6.86	<.01
Within (error)	194.50	90	2.16		

gested. The number of successes is 13 and 7, respectively. The next problem, Problem 6, it will be recalled, had a longer mean solution time than any of the others suggesting a higher difficulty level. At this point there is an abrupt drop in the number of successes among the low quartile groups. (It is also the point where failures first appeared in any number in the high groups.) Only four of the twenty who reached it completed it. It should be pointed out here that the average time available to these 20 *S*s when they started Problem 6 was 5.44—one minute longer than the average solution time for that problem. No low quartile *S* completed more than six problems. The proportion of successes—3 to 1—still favors the third quartile, but the numbers are so small that they can be only suggestive.

After considering these points, we suggest the explanation that the greater difficulty of Problem 6 prevented the more gradual dropping out of third quartile *S*s as compared to fourth quartile *S*s, since it stopped nearly all low *S*s. Had that problem been comparable to the others in level of difficulty, the trend started with Problem 5 might have continued more gradually. The difficulty level of Problem 6 also probably contributed to the lack of overlap in problem scores between the two halves of the motivation distribution. The tentative conclusion is that the difficulty of the task is one factor in determining the kind of relationship which will appear between achievement motivation and successful completion. This is a point for further investigation.

In summary, it may be said that the main hypothesis was clearly supported. It seems evident that, in this kind of a situation at least, performance depends not only on having the ability to do well but on the motivation to use that ability as well. The form the relation between achievement motivation and performance takes appears to depend in part on the characteristics of the task.

SUMMARY

Three hypotheses were tested: that AFQT would be related to problem-solving performance among *S*s with high achievement motivation but not among *S*s with low achievement motivation; that motivation would be related to performance when all *S*s worked the same length of time; and the relation between motivation and performance would be greater below the median of the motivation distribution than above it.

Ninety-six *S*s, eight in each of the 12 cells of a 3 × 4 table representing three levels of intelligence and four levels of motivation score, solved as many problems as they could in 25 minutes on a switch-light pattern apparatus.

The results indicated clear support for the first two hypotheses but none for the third. There is some indication that the difficulty of the task influences the motivation-performance relationship so that this will not be constant from task to task.

REFERENCES

1. FRENCH, ELIZABETH G. Some characteristics of achievement motivation. *J. exp. Psychol.*, 1955, **50**, 232–236.
2. FRENCH, ELIZABETH G. *Development of a measure of complex motivation. USAF Pers. Train. Res. Cent., Res. Rep.* April 1956, No. AFPTRC-TN-56-48.
3. FRENCH, ELIZABETH G., & THOMAS, F. H. The relation of achievement motivation to problem-solving effectiveness. *J. abnorm. soc. Psychol.*, 1958, **56**, 45–48.

Received August 13, 1957.

FAILURE AND THE SUBSEQUENT PERFORMANCE OF SCHIZOPHRENICS[1]

GORDON W. OLSON[2]

State University of Iowa

TO ACCOUNT for psychological deficit (9) in schizophrenia, the hypothesis is often advanced that schizophrenics are less responsive to environmental stimuli than are nonschizophrenics (5, 11). In direct contrast to this view, it has also been suggested that schizophrenics are withdrawn from social contacts because they are hypersensitive to social stimuli, notably censure (2). Supporting the hypothesis of hypersensitivity are the studies of Garmezy (6), Bleke (1), Webb (15), and Hahn (7). These investigators all reported deleterious effects on the performance of schizophrenics from social punishment. Cohen (3), on the other hand, found that a mild physical punishment (shock) was more successful in maintaining the response level of schizophrenics than a verbal "rapport" condition. These findings suggest a possibly important difference in the reaction of schizophrenics to mild social punishment and to mild physical punishment and lend weight to the idea of a specific hypersensitivity to social punishment in schizophrenia.

Studies of success and failure in normals typically yield results which appear to be directly opposite to those found for schizophrenics. Consistently, mild failure conditions appear to facilitate the performance of normal subjects to a greater extent than success conditions (4, 12, 13). Social rewards and social punishment seem, therefore, to produce directly opposite effects in schizophrenics and normals. In an attempt to clarify some of these issues, this study was designed to investigate the comparative effects of positive, negative, and nonevaluative verbal statements on the performance of schizophrenic and normal subjects.

METHOD

Task. The task was constructed with a view toward maximizing performance and minimizing new learning in an area involving verbal-symbolic performance. A digit symbol test was developed, requiring the substitution of a caret for the numbers 1, 2, 3, and 4, the apex of which rotated 90° clockwise for each consecutive number, i.e., 1 ∧, 2 >, 3 ∨, and 4 <. The measure used was the number of *attempted* items per 50-second trial. The number of attempted substitutions was used rather than the number correct because a pilot study revealed a correlation of +.95 for the two measures, and motivation, not efficiency, was the subject of main concern. The sequence of digits on the worksheet was derived from a table of random numbers, and when a number appeared twice in succession, the next occurring number replaced the duplicate. The substitution code appeared at the top of the worksheet. A maximum score of 100 was possible.

Subjects. The 45 schizophrenic *S*s were male patients at the Knoxville, Iowa, Veterans Administration Hospital between the ages of 20 and 40, with an educational achievement level of not less than the sixth grade. All patients carried an unchallenged diagnosis of schizophrenia which encompassed the usual subtypes. The length of hospitalization ranged from six months to 17 years, the mean being 4.3 years. The 45 normal *S*s were randomly selected male psychiatric aides at the same hospital who met the same age and educational criteria as the schizophrenics. No psychotic *S* had experienced psychosurgery or had received electric or insulin shock therapy within the previous two months. If *S* was presently on a tranquilizing drug, this fact was noted; the tranquilized *S*s seemed approximately equally distributed among the three instructional subgroups. A reasonable degree of homogeneity was sought by presenting each *S* with not more than two adaptation trials of 15 substitutions each prior to the experiment proper. Any *S* who failed to complete the second such trial with less than three errors was discontinued. Thus, the experimental *S*s were a select group of hospitalized schizophrenics with respect to cooperation and ability to learn a simple task.

Procedure. Prior to the first experimental trial, *S* learned the substitution code to two correct trials by orally reporting the direction of the apex when only the numeral was exposed on a card. The untimed adaptation trials of 15 items followed without reference to a cue sheet. Errors occurring during the adaptation trials were pointed out by *E* and corrected by *S* in order to further the learning process. Errorless performance on the first adaptation trial excused *S* from the second, and the experimental trials were immediately initiated.

For the experimental trials, the test sheets were drawn from a large pile within sight of *S*, who was not told how many trials were forthcoming. Each trial consisted of 50 seconds of work followed by a 10-second rest. Five such trials constituted a block, and the three blocks were separated by six-minute rest intervals during which *E* scored the preceding five trials. Positive,

[1] Based upon a doctoral dissertation submitted to the State University of Iowa. The author wishes to express his gratitude to the VA Hospital, Knoxville, Iowa, for the cooperation necessary for conducting this study.

[2] Now at Anoka State Hospital, Anoka, Minnesota.

negative, or nonevaluative conditions were introduced during the last minute of the rest intervals between Blocks I and II and Blocks II and III, but not between trials within blocks. Assignment to the subgroups defined by the conditions was consecutive.

The rest interval statements were as follows, with only minor modifications to insure naturalness and conviction; the emphasis was on instituting the desired state, rather than presenting an exact sequence of words:

Positive: I have scored your papers, and I see that so far you are doing very well. As a matter of fact, you are doing much better than most men of your age and background. Let's try it again and see what happens.

Negative: I have scored your papers, and I see that so far you are not doing very well. As a matter of fact, you are doing very poorly compared with other men of your age and background. Let's try it again and see what happens.

Nonevaluative: I have scored your papers during the rest period which we just had, but we still have some more to do, so give me just a second or two to get the material rearranged and then let's do it again and see what happens.

Through Block I only two groups are differentiated, schizophrenic (S) and control (C). For Blocks II and III, six groups were designated, identified by appropriate symbols: schizophrenic, nonevaluative (SO); schizophrenic, positive (SP); schizophrenic, negative (SN); control, nonevaluative (CO); control, positive (CP); control, negative (CN). There was no predominance of any subtype of schizophrenia in any of these categories. Reassurance of adequate performance was given both the SN and CN groups after completion of the experiment.

RESULTS

Among the patients tested, six were unable or unwilling to participate at the required criterion level. Among the normal *S*s, all scheduled for participation took part with the exception of one who requested exclusion for personal reasons. The mean ages for the S and C groups were 29.4 years and 33.2 years, respectively, and the standard deviations were 5.21 and 3.95. While the normal group was significantly older ($p < .01$) than the patient group, the negative correlation of age and performance should attentuate rather than exaggerate any differences favoring the normals. The mean educational level for the schizophrenics was 10.9; for the normals it was 10.4. The standard deviations were 2.75 and 1.89, respectively.

The mean numbers of attempted substitutions for each subgroup are presented in Table 1. Analysis of the results for Block I indicates that the initial performance of the schizophrenic group was clearly inferior to that of the normals ($p < .01$), although both groups showed significant improvement over the five trials ($p < .01$). Analysis of variance also indicates that both diagnostic groups and all six subgroups improved over these successive trials at the same rate, and that there was no significant difference in performance over the initial trials among the subgroups by conditions.

Employing a Lindquist (10) Type III design, the analysis of variance of the two diagnostic groups, the three subgroups by conditions, and the three blocks again indicates superior performance for the normal group ($p < .01$). Significant improvement in performance across blocks is also demonstrated ($p < .01$), as is significant improvement in performance due to conditions ($p < .05$). None of the interactive effects were significant. Analysis of the effects of the conditions for Blocks II and III reveals a higher level of performance for *S*s under either the positive or the negative statements than those under the nonevaluative condition ($p < .01$). The positive and negative conditions did not appear to differ in their effectiveness.

Analysis of the data for the schizophrenics

TABLE 1
MEAN NUMBER OF ATTEMPTED SUBSTITUTIONS PER TRIAL

Group	Trials														
						Block II					Block III				
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
CP	31.3	32.1	33.5	33.5	33.7	40.5	39.5	39.5	38.4	38.6	45.1	43.3	42.9	41.9	40.9
CN	28.0	30.7	32.1	32.5	32.6	40.6	43.3	44.3	44.3	43.6	48.6	48.9	48.1	48.5	48.2
CO	28.5	30.3	31.6	32.3	33.9	37.2	36.7	37.8	37.2	37.4	42.3	40.7	40.1	39.9	39.3
SP	24.2	26.6	26.9	28.3	29.3	34.6	35.5	36.7	36.8	37.0	41.7	40.1	40.5	40.7	40.1
SN	23.0	25.5	25.7	25.9	28.7	32.7	33.6	32.9	33.2	34.1	36.9	37.3	37.9	36.5	35.6
SO	23.7	25.3	27.3	27.7	29.1	31.7	30.7	30.9	30.2	30.9	35.0	34.5	32.2	32.7	30.5

only reveals an interaction between conditions and blocks ($p < .01$), indicating the differential influence of the statements on the subsequent performance of the schizophrenic *S*s. Both positive and negative conditions produced enhanced schizophrenic performance as compared with nonevaluation (in both cases, $p < .01$), and the positive statements resulted in significantly greater improvement than did the negative ($p < .01$).

Among the normals only, there was a significant interaction of instructions and blocks ($p < .01$). Study of the simple effects of conditions indicated that negative statements resulted in significantly improved performance as compared with the nonevaluative condition ($p < .01$), but the positive instructions apparently did not. Also, the negative condition for normal *S*s was followed by significantly better performance than was the case for positive statements ($p < .01$).

A comparison of the simple effects of conditions between diagnostic groups showed differential effects. In the two nonevaluative groups, the normals improved more over blocks than did the schizophrenics when neither positive nor negative statements were made ($p < .01$). The schizophrenics, however, performed better than the controls when both were under the positive condition ($p < .01$), and the controls showed significantly greater improvement than the schizophrenics when both were under the negative condition ($p < .01$).

It is of interest to examine briefly the relative amount of improvement shown by the subgroups as a function of rest and conditions. When improvement is defined as the amount of gain from the last prerest trials to the first postrest trial, there was an insignificant over-all difference in improvement scores in favor of the normals. The difference in improvement was significant ($p < .05$) when improvement after the first rest period was considered separately, but the effect was not present following the second rest period. Since the over-all improvement data was confounded with the effects of the conditions, comparison was made of the improvement scores for the two nonevaluative groups only. No difference in improvement scores after either rest was found.

DISCUSSION

The results of this study indicate that negative statements, while not harmful to the performance of schizophrenics, were less effective as a facilitator of performance than was praise; similarly, failure was a less effective enhancer of performance for schizophrenics than it was for normals. This interpretation is in marked contrast to the conclusions of other investigators (1, 6), who in similar studies underscored the detrimental effects of social punishment in schizophrenia. The point merits emphasis that the present conclusions are based on the comparisons under a nonevaluative condition, a condition not included in either of the studies by Garmezy or Bleke.

Consistency of results of the present study and the above studies is found in the differential performance of the normals and schizophrenics to the positive and negative conditions. In each study, the schizophrenic *S*s did less well under negative conditions than positive; and in both Garmezy's and the present study, the normals showed better performance under the negative statements than positive. Although the differential effect of praise and failure for the diagnostic groups falls short of significance in the present study ($p > .10$), this differential response appears in the increasing difference over trials under the negative condition and the decreasing difference over trials under the positive condition.

The theory involving hypersensitivity to failure in schizophrenia (1) calls for the obvious prediction of a greater gain in performance for the normal negative than for the schizophrenic negative group, and this prediction was borne out here. A second prediction from this theory is that the schizophrenic negative group would do less well than the schizophrenic positive group; this difference, too, was found. Thus, the hypothesis of the deleterious effect of social punishment *relative to social reward* in schizophrenia seems substantiated. That the deleterious effect of punishment is only *relative* to the positive condition is clearly shown by the significantly greater improvement of the schizophrenic negative group over the schizophrenic nonevaluative group, in that if social punishment were absolutely deleterious, it would be necessary that the schizophrenic

negative group do less well than the schizophrenic nonevaluative group.

It could be suggested that the initial deficit shown by the schizophrenics was due to a deficit in motivation, except that the significant difference in Block III still favored the normals. Thus, the initial diagnostic difference was not overcome by enhancement reflecting possible increments in motivation. Nevertheless, the enhanced performance of schizophrenics on this task equalled the unenhanced performance of normals, suggesting that motivational factors may have compensated for the diagnostic difference.

The finding of an increasing difference between the nonevaluative groups is interesting with respect to the concept of a differential rate of accumulation of inhibition for schizophrenics and normals (14). Support for such a notion is found in the increasing difference between the diagnostic groups under the nonevaluative condition, but refutation is found in the absence of differences in performance by the groups over trials within blocks. After this hypothesis, and following the paradigm of Hull (8) for reactive inhibition, it would be expected that the schizophrenic group would show more decline over the trials within blocks than would the normals. In addition, the scores for improvement after rest show that there was a greater amount of improvement for the normals, a finding in opposition to an interpretation of a greater accumulation of reactive inhibition for the schizophrenic *S*s.

SUMMARY

The present study investigated the questions of whether deficit would obtain for schizophrenics on a timed task of nonsocial content and, if so, whether the deficit would increase or decrease over trials under a condition of nonreinforcement and, finally, whether success or failure instructions would differentially affect the subsequent performance of schizophrenics and normals.

Forty-five schizophrenic males and 45 normal males were administered 15 timed trials of a simple digit-symbol task and the number of attempted substitutions per trial recorded. During the two rest periods which followed each block of five trials, the *S*s were consecutively assigned to positive, negative, and nonevaluative groups and informed respectively that they were doing very well, very poorly, or they were administered no evaluative comments relative to their immediately prior performance.

The results indicated that both schizophrenics and normals responded with enhanced performance to the negative conditions, but that only the schizophrenics demonstrated enhanced performance under the positive condition. There is suggestive evidence that the positive and negative statements affected schizophrenics and normals in a differential manner, with schizophrenics benefiting more from the positive than from the negative, but there is no evidence to suggest that the negative condition had a *detrimental* effect upon schizophrenics. The data are consistent with other studies reporting differential effects of social punishment and social reward for schizophrenics and normals. The data do not support the concept of increased reactive inhibition in schizophrenia, when the Hullian paradigm for reactive inhibition is employed.

REFERENCES

1. BLEKE, R. Reward and punishment as determiners of reminiscence effects in schizophrenic and normal subjects. *J. Pers.*, 1955, **23**, 479–498.
2. CAMERON, N. Reasoning, regression, and communication in schizophrenics. *Psychol. Monogr.*, 1938, **50**, 1–34.
3. COHEN, B. D. Motivation and performance in schizophrenia. *J. abnorm. soc. Psychol.*, 1956, **52**, 186–190.
4. FARBER, I. E., RUSSELL, W. A., & ANDREAS, B. G. Effect of failure upon performance in verbal and motor learning situations. *Amer. Psychologist*, 1949, **4**, 351. (Abstract)
5. FENICHEL, O. *The psychoanalytic theory of neuroses.* New York: Norton, 1945.
6. GARMEZY, N. Stimulus differentiation by schizophrenic and normal subjects under conditions of reward and punishment. *J. Pers.*, 1952, **20**, 253–276.
7. HAHN, I. N. The effects of success and failure experiences on chronic schizophrenic patients. Paper read at Amer. Psychol. Ass., Chicago, 1956.
8. HULL, C. L. *Principles of behavior.* New York: Appelton-Century, 1943.
9. HUNT, J. MCV., & COFER, C. N. Psychological deficit. In J. McV. Hunt (Ed.), *Personality and the behavior disorders.* Vol. II. New York: Ronald, 1944. Pp. 971–1032.

10. LINDQUIST, E. F. *The design and analysis of experiments in psychology and education.* Boston: Houghton Mifflin, 1953.
11. PIOTROWSKI, Z. Work habits as a means of differentiating native and acquired deficiency. *Psychol. Bull.*, 1933, **30**, 728–729.
12. RUSSELL, W. A., & FARBER, I. E. Retention of verbal material as a function of degree of failure experienced in original learning. *Amer. Psychologist*, 1948, **3**, 355. (Abstract)
13. STEISEL, I. M., & COHEN, B. D. The effects of two degrees of failure on level of aspiration and performance. *J. abnorm. soc. Psychol.*, 1951, **46**, 79–82.
14. VENABLES, P. H., & TIZARD, J. Performance of functional psychotics on a repetitive task. *J. abnorm. soc. Psychol.*, 1956, **53**, 23–26.
15. WEBB, W. W. Conceptual ability of schizophrenics as a function of threat of failure. *J. abnorm. soc. Psychol.*, 1955, **50**, 221–224.

Received August 16, 1957.

THE EFFECTS OF PLAUSIBILITY ON THE LEARNING OF CONTROVERSIAL STATEMENTS[1]

EDWARD E. JONES AND RIKA KOHLER
Duke University

PERHAPS one of the most ubiquitous assumptions of recent research on human cognition is that perception, learning, and memory operate in ways which are instrumental to adjustment and survival. This assumption has given rise to a number of experiments in the tradition of perception research which have shown, for example, that our cognitive activities tend to be selective (12, 19), that we accentuate (9) or fixate (17) stimuli that are symbolic of reward, and we fail to perceive (13) or autistically distort (15) stimuli that are inimical to our own values or self concepts. These phenomena all seem to exemplify the trend toward cognitive constancy or stability of attitude.

While many of these studies have been criticized on both procedural and logical grounds (e.g., 14, 16), there seems less resistance to the contention that the so-called higher mental processes of learning and association are conditioned by adjustive requirements. The most clear-cut demonstrations of such effects have occurred in the area of meaningful learning, where numerous attempts have been made to show that we learn best those statements which are congenial to our beliefs and attitudes (e.g., 4, 10, 18, 20). The modest but fairly consistent success of these attempts has given support to those impressed by the autistic nature of learning. The proposition that we learn best that with which we agree was originally derived from basic considerations of the functional or instrumental role of learning. The generality of this proposition depends, however, on the extent to which it really is adjustive to be especially retentive of congenial views.

The fact that we are, under normal circumstances, relatively unreceptive to facts and arguments that run counter to our biases would appear to be a potent barrier to attitude change. To the extent that autism is operative, the recipient of an influence communication will have difficulty learning the communication message and/or will distort the content of the message so that it conforms more closely to the position he already holds.

It is fortunate for the experimenter, the propagandist, or the educator who attempts to modify a significant attitude that all learning and retention is not autistic, at least in the more conventional sense of the word. In the first place, the very fact of exposure is certainly important. An individual may retain more of a co-valuant than a contra-valuant communication, but he still may retain significant portions of the latter. Furthermore, it is quite easy to create conditions in which a reversal of the simple autism effect occurs. Jones and Aneshansel (8) have shown how prosegregation *S*s learn antisegregation statements *better* than antisegregation *S*s if both groups anticipate a task of providing appropriate counter-arguments to prosegregation statements. When there is no mention of any counter-argument task, the usual autism effect demonstrated most clearly by Levine and Murphy (10) is observed. Thus, *S*s may be motivated by an ulterior incentive to learn well those statements with which they disagree, a finding which bears a relation to studies of role-playing in attitude change research (cf. 7).

Other considerations have to do with the plausibility or extremity of the arguments involved in the communication. It seems reasonable to assume, as Hovland, Janis, and Kelley (5) have done, that "a communication only slightly removed from the individual's position will sometimes influence him positively while one far removed will affect him negatively (boomerang effect)"[2] (p. 286). What is the significance of this assumption for the learning process itself? Obviously, the extreme

[1] This study was conducted in the program of the Organization Research Group of the Institute for Research in Social Science at the University of North Carolina. The O.R.G. is supported jointly by the Ford Foundation and the Office of Naval Research [Nonr-855 (04)]. The research was partially financed by a grant to the senior author from the Duke University Research Council.

[2] A recent study by Hovland and Pritzker (6) indicates that this does *not* happen when the issue is relatively uninvolving and the prestige of the source is high.

message must be assimilated in some way in order for a boomerang effect to occur. Beyond this, however, the individual may be positively motivated to absorb and fixate the extreme statement. He may be avidly receptive to ludicrous overstatements of positions differing from his own, just as the arch conservative tends to fixate in his memory statements made by liberal associates which may be construed as "socialistic." One's own self-righteousness is enhanced by indications of the implausibility of the opposition; therefore, such indications will be retained and cherished as self-supportive.

The converse is also true. We are embarrassed when others give ineffective support to our own cause, just as the educated Southerner may be embarrassed by the Georgia cracker. As a consequence, we rush to forget the false and implausible arguments that favor our position but make salient the negative elements in our own reference group. If these common sense deductions are true, then we would have to qualify further the statement that people learn best those statements which are in favor of their position. In order for the conventional autism effect to hold, the statement would have to be sufficiently plausible and effective so that the individual could comfortably endorse both the position implied by the statement and the supporting reasons adduced.

The present experiment is concerned with the effects of both plausibility and direction on the learning of controversial statements. In brief outline, we propose to examine the proposition that *S*s who hold to one set of beliefs learn plausible statements in favor of those beliefs and implausible statements opposing them better than they learn implausible-favoring or plausible-opposing statements. In more specific terms, three related hypotheses are tested:

1. Prosegregation *S*s learn plausible prosegregation and implausible antisegregation statements with less difficulty than they learn plausible antisegregation and implausible prosegregation statements.

2. The reverse is true of the antisegregation *S*s.

3. Those with neutral or intermediate beliefs learn all types of statements equally well.

A preliminary experiment and a more rigorous replication experiment were conducted to test these hypotheses.

EXPERIMENT I

Method

Subjects. A questionnaire obviously dealing with attitudes toward segregation was administered to 299 introductory psychology students. One set of items on this questionnaire constituted a self-rating scale which the students could check from "1. Strong prosegregation. Example: I would favor all measures short of actual violence to insure the long-run maintenance of segregation" to "9. Strong prointegration. Example: I would favor relying on the full power of Federal and State authorities to insure that desegregation occurs as rapidly as possible." After indicating their position "in general" on this nine-point scale, the students were confronted with a number of community situations that would presumably make integration difficult to effect (e.g., "There are many more Negro school-age children than white, and almost as many Negro teachers as white teachers"). For each situation, the students were to check a number from one to nine, indicating their action attitude as a responsible member of the community described. Their initial self-rating was averaged with self-ratings repeated across five community situations to obtain a final attitude score. From the pool of 299 students, 45 *S*s were selected: 15 prosegregation *S*s (hereafter referred to as "pros") with scores ranging from 1.11 to 2.50, 15 intermediate *S*s (hereafter referred to as "neutrals") with scores ranging from 3.50 to 4.50, and 15 antisegregation *S*s (hereafter called "antis") with scores ranging from 6.85 to 7.98.

Experimental procedure. The actual procedure involved in the learning task was very similar to that reported in (8). Each of the 45 *S*s attempted to learn 12 statements in individual sessions. The statements were printed on separate cards and each *S* was instructed to read the cards aloud at normal reading speed. After reading the 12 statements in this manner, *S* was asked to reproduce as many statements or phrases as he could remember. This identical procedure was repeated for five learning trials. The order of statements differed randomly on each trial, but the order was constant across *S*s.

Six of the 12 statements to be learned were prosegregation in direction, six were antisegregation. Within each group of six statements, three were "plausible" and three were "implausible." Plausibility was determined by the pooled rankings of eight graduate student judges. Each judge was presented with 20 statements, 10 of which were obviously prosegregation and 10 of which were obviously antisegregation. Out of each group of 10 statements, the three which were ranked lowest and the three which were ranked highest on "plausibility" were selected for use in the experiment. All statements were equated for length. Examples of each type of statement follow:

1. *Plausible prosegregation.* Southerners will have to pay the price of lowered scholastic standards if they yield to the pressures to integrate their schools.

2. *Implausible prosegregation.* If Negroes and whites were meant to live together, they never would have been separated at the beginning of history.

3. *Plausible antisegregation.* The present inferior condition of the Negro is the result of long and effective suppression by the Southern whites.

4. *Implausible antisegregation.* The real reason why most Southern whites oppose integregation is their realization that the Negro is more capable than they are.

While there is room for debate concerning the sense in which a given statement is plausible or implausible, it was important to avoid statements which were completely ridiculous as well as those which were completely truistic. The judges, thinking in terms of the experimental *S*s that were to be used, showed a high degree of agreement in their rankings of statements.

Results

Each reproduction attempt was given a score of 1 if it expressed a vague idea of part but not all of the statement; 2 if it expressed the over-all idea of the entire statement; 3 if it was correct except for the use of synonyms; and 4 if it was verbatim. Because of the greater possibility on the earlier trials of reproductions that combined statements or fell between statements, and the consequent unreliability of scoring, analysis was restricted to reproductions made on Trial 5.

The results of Experiment I provide rather firm support for the hypotheses. As Table 1 indicates, the appropriate comparisons between pro and anti *S*s on each of the four statement types are in the right direction in every case, and three of the four differences are significant at the .05 level or better. In every case but one, the neutral *S*s fall between the pro and the anti *S*s on their fifth trial learning scores. It would appear that plausibility, in combination with the direction of a controversial statement, is a variable of some importance in predicting learning success. The main reversal of prediction is accounted for by the antis' relatively poor learning of the plausible anti statements and their relatively good learning of the implausible anti statements.

In spite of the generally good support for the experimental hypotheses, there are a number of procedural criticisms that might be directed at this first experiment. The reliability and validity of the attitude self-ratings are unknown, *E* may have unwittingly influenced the results by on-the-spot scoring, the sensitivity and reliability of scoring procedures which consider only the final trial might be considered suspect, and the generalization value of the results is contingent on the representativeness of the particular statements chosen as plausible and implausible. For these and other reasons a replication was designed.

TABLE 1
MEAN LEARNING SCORES FOR TRIAL 5 BY ATTITUDE AND STATEMENT TYPE
(Experiment I)

Subjects	Statements			
	Pro		Anti	
	Plausible	Implausible	Plausible	Implausible
Pro	2.7	1.4	1.3	2.6
Neutral	2.1	2.0	2.4	2.3
Anti	1.5	2.6	2.2	2.1
Anti vs. Pro P_{diff}[a]	$<.01$	$<.01$	$<.05$	—

[a] By *t* test, two-tailed.

EXPERIMENT II

Method

The learning procedure itself was identical to that of the first experiment except that *E* did not know the *S*'s attitude category until the session was over, and each reproduction was transcribed verbatim on specially designed recording sheets. The major changes occurred in the following respects.

Subjects. Once again introductory psychology students served as *S*s, but they were selected by a different attitude measure. A 10-item Likert-type scale was administered to 147 students. This scale was identical to that used by Jones and Aneshansel (8), and its characteristics and validity are described in the earlier article. In the present study, scores ranged from a low of 13 (possible minimum, 10) to a high of 66 (possible maximum, 70) and the distribution was somewhat skewed toward the low, or antisegregation, end of the scale. Thirty-nine *S*s were selected: 13 pros with scores ranging from 43 to 66, 13 neutrals with scores from 28 to 34, and 13 antis with scores from 13 to 21.

Statements. Twelve statements, all of which differed from those used in the first experiment, were used in the learning task. Once again, three statements were pro-plausible, three anti-plausible, three pro-implausible, and three anti-implausible. The same procedure was followed to determine statement plausibility, namely the pooled rankings of, this time, six judges.

Scoring. Each statement was divided into four items, which represented single idea units. An item was then scored as correct if it was reproduced verbatim or only incidentally modified by the use of synonyms. As a check on the reliability of scoring, 1459 items were systematically selected to sample equally from the 39 *S*s and the 12 statements. These items were independently scored by a second rater and there was 90% agreement on categorization of "correct" items.

Results

The results of the second experiment are even more dramatic than those of the first. The findings are graphically summarized in Fig. 1. Here it may be seen that the hypotheses are confirmed in every detail, and for each type of

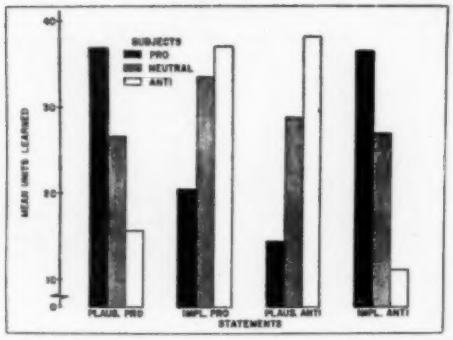

FIG. 1. THE RELATION BETWEEN PLAUSIBILITY, DIRECTION OF STATEMENT, AND ATTITUDE IN THE LEARNING OF CONTROVERSIAL MATERIAL

statement the neutral *S*s fall between the pros and the antis. Thus the familiar Levine and Murphy effect holds only for plausible statements. As far as implausible statements are concerned, *S*s tend to learn better those statements which are opposite in direction to their own position.

The major statistical analysis of these results involved assigning each *S* a score representing the number of items correctly reproduced for each type of statement on each of the five learning trials. In order first to determine the effect of trials, these scores were placed in a mixed factorial design similar to that described by Edwards (3) for repeated measures on the same *S*s. Using the pooled within-*S*s error term advocated by Edwards, the main effect of trials was highly significant ($p < .001$), but none of the interactions *involving trials* approached significant values. Therefore, items learned to the criterion were summed over all five trials for each *S*, and the resulting scores were analyzed in a Type VI mixed design (following Lindquist, [11]). Collapsing over trials in this manner made it possible to use more conservative, and more precisely appropriate, error terms for each statistical effect without making the design impossibly cumbersome.

The results of this major analysis are presented in Table 2. In statistical terms, our hypotheses combine into the single prediction of a significant interaction between attitude, statement direction (pro or anti), and plausibility. With 2 and 36 degrees of freedom, an *F* ratio of 8.42 would be significant at the .001

TABLE 2

ANALYSIS OF VARIANCE OF LEARNING SCORES POOLED OVER FIVE TRIALS
(Experiment II)

Source	*df*	Mean Square	Error Term	*F*
A: Attitude	2	105.63	b	<1.00
Error b	36	274.03		
Between *S*s	38			
B: Direction	1	210.01	W_1	5.18*
C: Plausibility	1	16.68	W_2	<1.00
A × B	2	22.51	W_1	<1.00
A × C	2	148.86	W_2	6.17**
B × C	1	354.00	W_3	3.24
A × B × C	2	5902.63	W_3	53.99***
Error W	108	58.00		
Error W_1	36	40.53		
Error W_2	36	24.13		
Error W_3	36	109.33		
Within *S*s	117			
Total	155			

* $p < .05$.
** $p < .01$.
*** $p < .001$.

level. The obtained *F* ratio for this interaction of 53.99 is far beyond the range of conventionally tabled probability values and is clearly significant in the extreme.

There are two other significant sources of variation, ignoring the above mentioned main effect for trials which, as plotting the means shows, is a function of the increase in number of items correctly reported from trial to trial. The significant main effect of direction reflects the fact that slightly more pro than anti statements are learned overall. The significant interaction between attitude and plausibility tells us that in the present experiment antis tended to learn plausible statements better, while pros tended to learn implausible statements better.

In view of the design of the experiment and the fact that each of these significant effects is dwarfed in magnitude by the higher-order interaction, the generalization value of these unpredicted differences would probably be limited in any event. In addition, none of these significant effects (except that of trials and the predicted higher-order interaction) replicates findings of the first experiment. In fact, more anti than pro statements were learned overall

TABLE 3
COMPARISONS BETWEEN ATTITUDE GROUPS FOR DIFFERENT STATEMENT TYPES
(Probability values for t test results, Experiment II)

Statement Type							
Pro Plausible		Pro Implausible		Anti Plausible		Anti Implausible	
Direction	P	Direction	P	Direction	P	Direction	P
P > N	.05	A > N	.50	A > N	.10	P > N	.05
N > A	.10	N > P	.01	N > P	.01	N > A	.001
P > A	.001	A > P	.01	A > P	.001	P > A	.001

Note.—P, N, and A stand for pro *Ss*, neutral *Ss*, and anti *Ss* respectively.

in that experiment. In addition, the anti *Ss* actually learned more implausible and fewer plausible statements than did the pro *Ss*.

It is important to note that the *S*'s attitude bears no relation to the total amount of learning that occurred. This finding removes any reason for concern with the differential intelligence or learning ability of pro and anti *Ss* in the present experiment.

While the analysis of variance is extremely useful in correcting for the lack of independence between the probability values of individual paired comparisons, the interaction between attitude, direction, and plausibility is impossible to interpret without reference to Fig. 1. To facilitate interpretation further, probability values of t-test comparisons between the means of different subject groups are presented in Table 3. From this table it becomes clear that the antis not only differ from the pros at the .01 level for every comparison; antis and pros in turn differ significantly from the neutral *Ss* on five out of eight comparisons.

DISCUSSION

Both experiments provide strong support for the experimental hypotheses, which concern the importance of plausibility in the learning of controversial statements. When *S* is confronted with plausible statements that could be effectively used either in support of or in opposition to his own position, he tends to learn best those statements which argue in the direction of his own attitudinal bias. When the statements are overdrawn, implausible, or based on fallacious premises, however, *S* tends to learn best those statements which argue *against* his own attitudinal bias. Neither plausibility per se, nor the direction of the statement per se, seems to affect the learning process (cf. 20).

Those who favor a functionalistic view of cognitive processes will have little difficulty assimilating the results of the present experiment. Working from the assumption that *S* is concerned with protecting his attitudinal position, we see him ignoring or responding defensively to the effective or plausible opposing argument and reacting with almost hypervigilance to the implausible opposing argument. It is usually assumed that the greater the threat to the organism, the more likely will vigilance predominate over defense (e.g., 2). This assumption would imply that the implausible opposing argument is more threatening than the plausible, an unlikely state of affairs in any context simulating a debate in which rationality is rewarded. Furthermore, such an assumption leaves it very difficult to explain why *Ss* have so much difficulty in learning the implausible statement that favors their own position. The explanation that we prefer is that the protecting of an attitudinal position is furthered by the denial both of implausible elements in one's own belief and plausible elements in opposing beliefs. Correspondingly, there is a tendency to fixate plausible favoring and implausible opposing beliefs.

The magnitude of the present effect is so striking that it prompts consideration of the possibilities of such a learning task for application to the assessment of attitudes. Particularly when one is dealing with attitudes that are opposed to cultural values and susceptible to dissimulation, it may be worth while to consider the joint manipulation of plausibility and direction in a learning task. Looking once again at the data of the second experiment, if we devise an index for each *S* that represents the difference between pro plausible and anti implausible on the one hand and anti plausible and pro implausible on the other, there is not one case of overlapping scores for pro and anti *Ss* as determined by the Jones-Aneshansel scale. Every score for the antis is less than 1, and all but one of the pros' scores is greater than 1. Whether such a high degree of discrimination would be found using other attitude measures with learning tasks of relevant materials is a matter for further empirical study.

SUMMARY

Two closely related experiments were conducted to test the following hypotheses: (*a*) prosegregation *S*s learn plausible prosegregation and implausible antisegregation statements with less difficulty than they learn plausible antisegregation and implausible prosegregation statements; (*b*) the reverse is true of the antisegregation *S*s; and (*c*) those with neutral or intermediate beliefs learn all types of statements equally well. Each hypothesis was confirmed, and the reliability of the findings was especially striking in the replication experiment where the amount learned was more carefully scored, and other precautions were taken against biasing the results.

While the results show that learning of a controversial statement cannot be predicted solely from the direction of the argument, they still support the functionalist's basic assumption that cognitive processes operate so as to promote the constancy of attitude and belief.

REFERENCES

1. ALPER, THELMA G., & KORCHIN, S. J. Memory for socially relevant material. *J. abnorm. soc. Psychol.*, 1952, **47**, 25–37.
2. BRUNER, J. S., & POSTMAN, L. Emotional selectivity in perception and reaction. *J. Pers.*, 1947, **16**, 69–77.
3. EDWARDS, A. L. *Experimental design in psychological research.* New York: Rinehart, 1950.
4. EDWARDS, A. L. Political frames of reference as a factor influencing recognition. *J. abnorm. soc. Psychol.*, 1941, **36**, 34–50.
5. HOVLAND, C. I., JANIS, I. L., & KELLEY, H. H. *Communication and persuasion, psychological studies of opinion change.* New Haven: Yale Univer. Press, 1953.
6. HOVLAND, C. I., & PRITZKER, H. A. Extent of opinion change as a function of amount of change advocated. *J. abnorm. soc. Psychol.*, 1957, **54**, 257–262.
7. JANIS, I. L., & KING, B. T. The influence of role-playing on opinion change. *J. abnorm. soc. Psychol.*, 1954, **49**, 211–218.
8. JONES, E. E., & ANESHANSEL, JANE. The learning and utilization of contravaluant material. *J. abnorm. soc. Psychol.*, 1956, **53**, 27–33.
9. LAMBERT, W. W., SOLOMON, R. L., & WATSON, P. D. Reinforcement and extinction as factors in size estimation. *J. exp. Psychol.*, 1949, **39**, 637–641.
10. LEVINE, J. M., & MURPHY, G. The learning and forgetting of controversial material. *J. abnorm. soc. Psychol.*, 1943, **38**, 507–515.
11. LINDQUIST, E. F. *Design and analysis of experiments in psychology and education.* Boston: Houghton Mifflin, 1953.
12. MCCLELLAND, D. C., & LIBERMAN, A. M. The effect of need for achievement on recognition of need-related words. *J. Pers.*, 1949, **18**, 236–251.
13. MCGINNIES, E. Emotionality and perceptual defense. *Psychol. Rev.*, 1949, **56**, 244–251.
14. POSTMAN, L. The experimental analysis of motivational factors in perception. In J. S. Brown et al., *Current theory and research in motivation.* Lincoln: Univer. Nebraska Press, 1953.
15. POSTMAN, L., BRUNER, J. S., & MCGINNIES, E. Personal values as selective factors in perception. *J. abnorm. soc. Psychol.*, 1948, **43**, 142–154.
16. PRATT, C. C. The role of past experience in visual perception. *J. Psychol.*, 1950, **30**, 85–107.
17. SCHAFER, R., & MURPHY, G. The role of autism in visual figure-ground relationship. *J. exp. Psychol.*, 1943, **32**, 335–343.
18. TAFT, R. Selective recall and memory distortion of favorable and unfavorable material. *J. abnorm. soc. Psychol.*, 1954, **49**, 23–28.
19. VANDERPLAS, J. M., AND BLAKE, R. R. Selective sensitization in auditory perception. *J. Pers.*, 1949, **18**, 252–266.
20. WATSON, W. S., & HARTMAN, G. W. The rigidity of a basic attitudinal frame. *J. abnorm. soc. Psychol.*, 1939, **34**, 314–335.

Received August 16, 1957.

VICARIOUS EXPERIENCE AND PARTIAL REINFORCEMENT[1]

DONALD J. LEWIS[2] AND CARL P. DUNCAN

Northwestern University

PSYCHOLOGISTS and others have used a whole cluster of concepts when referring to one person, in some fashion, "taking the part of another." Among these concepts are "imitation," "copying," "identification," "the significant other," "the normative generalized other," "role," "role playing," etc. These terms refer to different manners and different degrees of vicarious functioning, going, along one dimension, from a rather mechanical following of another's movements to the "internalization" of another's value systems and emotions so that one person can almost live for the other.

The theoretical history of various forms of vicarious functioning is a long one, but the history of experimentation on the problem is very brief. One of the first theoretical and experimental approaches to the problem from a behavioral science point of view was that of Miller and Dollard (7), who distinguished between "imitation" and "copying." Imitation was the matching of subjects' (*Ss*') responses to those of a human model. Because the model's presence was a necessary condition for imitation, this kind of behavior was characterized as matched-dependent. "Copying" referred to *S*'s following the model's behavior, or his supposed behavior, partly on the basis of cues produced by the copier's own responses, even when the model was absent.

Osgood (9) and Lazowick (3) indicate that the basis of copying and identification (putting oneself in the place of another) is a mediational response (r_M) common to both model and *S*. A mediational response is behavior, frequently internal or unobserved, which intervenes between environmental stimuli and instrumental responses. A simple diagrammatic representation of a mediational response is indicated by the following: $S \rightarrow r_M \rightarrow s_M \rightarrow R$. "S" represents the external, environmental stimuli, and "r_M" represents an internal, often autonomic, response to S. Because r_M is a response, it therefore has the properties of all responses, one of which is the emission of stimuli. The stimuli emitted by r_M are symbolized as "s_M" and these lead directly to and are an important component of the total stimulus situation, which also includes S, evoking the final instrumental response R.

The concept of mediational response has been highly developed in the area of learning by Hull (1), Spence (11), and others. In Spence's theory, the theoretical concept K is regarded as representing the motivational properties of the r_M-s_M mechanism and is defined in terms of specific empirical variables, viz., the magnitude of reward and the delay of reward. The mediational response (fractional anticipatory goal response) has served S-R theory in some of the ways that "expectancy" has served Tolman and his followers. Spence, for example, has shown how latent learning can be derived with the help of r_M-s_M. Osgood (8) has indicated that *meaning* probably has its source in a mediating response, and that when an environmental S means the same thing to two different persons it is because the S evokes the same or similar r_Ms. Amsel (12) has used r_M to interpret some of the data of partial reinforcement, and the concept has also been used to predict new data as illustrated by the study of Kendler, Karasik, and Schrier (2).

The present study represents an attempt to determine some of the conditions under which one can benefit from the experience of others and to relate these conditions to S-R theory, using a mediational response mechanism.

Previous research (4) has shown that partial reinforcement effects can be obtained from humans *Ss* in a lever-pulling ("slot-machine") situation; those *Ss* who win all of the time (continuously) during an acquisition series quit more quickly, when given the opportunity, than do *Ss* who have won only part of the time and irregularly (partially). This study attempts to determine if these partial reinforcement effects can be obtained from *Ss* whose acquisition experience in the situation is purely vicarious, resulting either from watch-

[1] This study was supported by a grant to the authors from the National Science Foundation.

[2] Now at Louisiana State University, Baton Rouge, Louisiana.

ing someone else play or from simply hearing about the experiences of someone else who has played. The theoretical interpretation of what happens during this vicarious experience follows most closely that of Osgood (9) and Amsel (12).

Consider that *E*, acting as a model, plays the slot machine and that *S* watches him play. The external stimulus situation S is therefore very similar for both. Assume further that both *S* and *E* like to win and dislike to lose. In addition, they both know how to put coins in slots and pull levers, and these simple instrumental responses, from much past experience, are under the command of verbal stimuli. That is to say, all adult *S*s can follow, completely without experience in the present specific situation, the verbal instructions, "Put a coin in the slot and pull the lever."

All *S*s have in common, then: (*a*) the same external situation; (*b*) the same highly developed skill at the instrumental response of lever pulling; (*c*) the same verbal command over this response, and (*d*) the same emotional responses, positive for winning and negative for losing. The external situation is S, the instrumental response is R, and both the emotional responses and the verbal commands are r_Ms.

With this much in common for all *S*s, it is necessary to note the differences that develop as a function of the experimental variables. These differences will be most clear if we consider first only the person who actually plays the machine, *E* in this case. When *E*, under a continuous reinforcement schedule, pulls the lever and wins, a pleasurable (at least distinctive) emotional response results. This mediating response, resulting from reward, will be symbolized as r_R. The r_R yields stimuli—s_R—which, among others, lead to the instrumental response of the next lever pull and thus s_R becomes attached to lever pulling.[3] When the extinction series starts and lever-pulling responses are no longer rewarded, a different r_M occurs, one that has a displeasurable or frustration component—r_F. The occurrence of r_F and its associated stimuli s_F results in, among other things, a considerable change in stimuli during extinction from those that had been present during acquisition, and extinction proceeds rapidly.

[3] Assuming a continuity theory of reinforcement.

For the partially reinforced *E*s, the situation is somewhat different. Both winning and losing occur during acquisition and thus both r_R-s_R and r_F-s_F occur during acquisition. Since both s_R and s_F are followed by lever pulling, the instrumental response becomes attached to both the pleasurable and unpleasurable responses, and quitting is delayed when the extinction series starts.[4]

The theoretical interpretation is thus clear enough for *E* who actually plays the machine; it is not quite so clear for the *S*s who simply watch. Although it is not difficult to see that r_R and r_F occur in these *S*s, it is the connection of r_R-s_R and r_F-s_F to the lever-pulling response that presents a problem because *S* never actually pulls the lever. The solution lies in the fact that the instrumental lever-pulling response is strong in the response repertoire of both *S* and *E*, and this instrumental response is led to by such subvocal responses (and their associated stimuli) as, "Good! A win. Now it is time for the next trial. Pull the lever." These verbal mediational responses have been associated with lever pulling through much past experience.

If we assume that the external stimulus situation, as a result of the events of the previous trials, evokes the appropriate r_Ms (r_R or r_F) in both *S* and *E* and that the r_M leads to, and therefore gets associated with, the subvocalizing about the next trial, and if the subvocalizing is already hooked to the instrumental response, then a definite chain is established between stimulus and instrumental response in the *S*s who simply watch as well as in the *E*s who actually play. A diagram of the chain would look like this:

$$S \rightarrow r_M \rightarrow s_M \rightarrow \text{verbalization} \rightarrow R.$$

The verbalization → R part of the chain has been learned previous to the experiment by *S*, but the S → r_M link, and the s_M → verbalization link are learned here during the acquisition series. The series of events occurring during the experiment can be said to have the same meaning for both *S* and *E* because the same r_Ms occur.

[4] There are undoubtedly other reasons for the superior resistance to extinction following partial reinforcement. One is the differential change in orienting or observing responses (13) that take place as a result of these two kinds of reinforcement.

The preceding analysis holds for those *S*s who watch someone else play the machine. For those who simply hear a description of the play, the analysis is only slightly different. The external situation for them consists primarily of verbal stimuli; the description of what happens. These verbal stimuli, it is assumed, will also be able to evoke the relevant r_Ms. Thus a diagram of the chain of events is this: verbalization $\rightarrow r_M \rightarrow s_M \rightarrow$ verbalization $\rightarrow$ R.

To bring data to bear on these theoretical notions and on an important social psychological problem, a ten-group experiment was performed. Two of the groups were traditional partial reinforcement groups with 25% and 100% being the percentages of reward involved. Two groups (one with 25% and one with 100% reward) watched *E* play the acquisition series and won on the plays that *E* did. It was believed that actually winning would increase *S*s' degree of participation in the situation. Two more groups (25% and 100%) watched *E* play the acquisition series but did not participate in the winning. Two groups had a hypothetical acquisition series explained to them and won when the *E* hypothetically won. A final two groups had a hypothetical acquisition series explained to them, but the *S*s did not participate in the winning. After the acquisition series was actually or vicariously completed, *S* took over the machine and was told he could play as long as he wanted. Two response measures were taken. The primary one, and the one to which the theoretical treatment refers, is the number of plays to quitting. Also a measure of *S*'s expectation on each trial was taken.

METHOD

Apparatus

The apparatus has been described in detail elsewhere (5). It is a modified slot machine. The player puts a metal disk in a slot at the top and then pulls a lever. A whirring sound ensues, lights flash, and a disk either comes out of a payoff tray or does not. The player has access to a large supply of these disks and can play as long as he desires. There is a counter inside the machine which cumulates the total number of plays. The machine pays off on particular plays which are determined in advance by the *E*; payoffs do not depend on anything *S* does.

Design

Five pairs of groups were used, with one group in each pair having a 25% reward acquisition series and the other a 100% reward acquisition series. For one pair, the Control, *S* played the acquisition series himself, and, of course, went right on to extinction. For another pair, Watch and Win, *S* watched *E* play the acquisition series. Whenever *E* received a payoff disk, *S* was given one too. A third pair, Watch Only, differed from the Watch and Win pair only in that *S*s never received a payoff. A fourth pair, Explain and Win, was told orally about a hypothetical person who had just previously played the machine. The *S* was given a play-by-play description of the hypothetical person's acquisition series. When the hypothetical *E* won, *S* was given a disk too. The final pair was the Explain Only. The same explanation was given to this group as to the Explain and Win, but now the *S*s were not given disks at any time. The acquisition series consisted of the first eight trials, with the 100% *S*s of each pair being paid on each of the eight trials and the 25% *S*s being paid on only Trials 2 and 8. After the eighth trial, *S* played the machine until he said that he wanted to quit, but he never won again. He was told that all disks that he had in his possession could be turned in when he was finished for 10¢ each. This exchange, of course, actually took place. No *S* in any group knew that the machine would never pay off (extinction) after the eighth trial, no matter whether the first eight trials had been played by *S*, *E*, or some hypothetical third person.

The *S*s were also asked to state on each play, just before they pulled the lever, whether they expected to win or not. Secured to the front of the machine was a line-rating scale with the numbers 1 through 6 placed at equal intervals on the line. Above these numbers was the title, "Expectation of winning or not winning." Corresponding to the numerical scale points from 1 through 6 were the descriptions: "fairly sure of not winning," "mildly sure of not winning," "slightly sure of not winning," "slightly sure of winning," "mildly sure of winning," and "fairly sure of winning." On each trial, before he pulled the lever, *S* called out the numerical value corresponding to his level of expectancy. The *E* sat to one side of the room and recorded the expectancy stated on each trial by *S*.

Subjects

The *S*s were drawn from introductory psychology courses at Northwestern University and were assigned to the various experimental groups as they appeared in the laboratory. Twenty *S*s were assigned to each of the 10 experimental conditions, making a total of 200.

RESULTS

Mean Log Plays to Extinction

The total number of plays for each *S* was converted into common logarithms and the mean for each group determined. A simple *t* was computed between the 100% and 25% mean log plays for the Control, and the usual partial reinforcement effect was found ($t = 2.36$ for 38 *df*). Then a succession of four 2×2 factorial analyses of variance was computed,

comparing each of the experimental groups against the control. One such comparison, for example, had 25% and 100% for the columns, and the Control condition against the Watch and Win condition for the rows. Another comparison had 25% and 100% for the columns, and the Control condition against Explain and Win for the rows.

The comparisons of Watch and Win against Control showed no difference. Vicarious experience, in other words, was as effective as the actual playing. The Watch and Win group actually played longer, but insignificantly so, than the Control. The percentage of reinforcement was significant ($F = 4.90$ for 1 and 76 *df*), with more plays to extinction for the 25% groups (see Fig. 1) and the interaction was not ($F = 2.54$ for 1 and 76 *df*). Thus the partial effect occurred in both the vicarious groups and the actually playing group, and not differentially.

The comparison of Explain and Win against Control also showed no significant difference between the actual and the vicarious experience groups. The percentage of reinforcement was significant at the .01 level ($F = 7.33$ for 1 and 76 *df*), and again there was no significant interaction ($F = 2.05$ for 1 and 76 *df*).

The comparisons of Watch Only against Control, and Explain Only against Control gave no significant main effects, but in both cases there was a significant interaction at the .05 level. This interaction is due to the effect that percentage of reward had on the control group but not on the experimental groups.

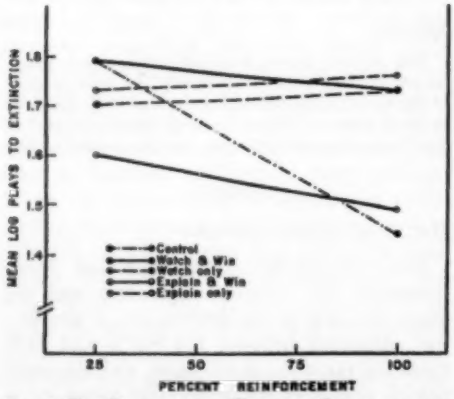

FIG. 1. THE RELATIONSHIP BETWEEN PERCENTAGE OF REINFORCEMENT AND MEAN LOG PLAYS TO EXTINCTION

As far as mean log plays to extinction are concerned, then, the vicarious groups did as well as the control groups overall. The partial reinforcement effect, however, appeared only in those situations where the *S*s actually won during the acquisition series.

Expectancies During Extinction

Because *S*s played different numbers of extinction plays, the extinction trials were Vincentized into ten blocks of trials for each *S*. The expectancy score for each *S* on the ten extinction blocks consists of the sum of his expectancies scores over that block of trials divided by the number of trials in the block. Therefore, the score for any block can be any value from one to six, the same values as in the original scale. Any score over three represents an expectancy that the machine will pay off on the next trial; a score below three represents an expectancy that the machine will not pay off on the next trial.

A three-point "trial-by-trial" trend analysis was performed on the data with the mean of the first two Vincentized tenths, the mean of the fifth and sixth tenths, and the mean of the last two tenths serving as the three points for the trend. Again a succession of 2×2 analyses of variance compared each of the experimental groups against the control for the trend. There were no significant main effects for any of the comparisons, but several interactive trend effects emerged as significant.

For the comparisons concerning Watch and Win vs. Control, and Explain and Win vs. Control, the trials by percentage interaction were significant at the .05 level. A graph of these data (not presented here) shows that the expectancy curve after 100% reinforcement falls off faster and ends up lower than does the curve after 50% reinforcement. For the comparisons with respect to Watch Only vs. Control, and Explain Only vs. Control, percentage of reinforcement was not an effective variable on trend of expectancies. In order for there to be a differential change of expectancies during extinction it was necessary for the *S*s actually to win during the acquisition series.

For all four of the 2×2 comparisons there was a significant triple interaction, trials by percentage by experimental condition. In each case, this is interpreted to mean that percentage of reinforcement had a greater effect

upon the control group expectancies than upon the experimental group expectancies.

DISCUSSION

The evidence indicates that the experimental variables had similar effects upon both response measures—mean log plays to extinction, and extinction expectancies; there were no over-all differences between any of the vicarious groups and the controls, e.g., vicarious and control *S*s showed the same over-all number of plays to extinction. The partial reinforcement effect—a greater number of plays to extinction for the 25% *S*s over the 100% *S*s—occurred, however, only in those vicarious groups which had participated in the winning. Previous research (5, 6) has revealed a characteristic pattern of extinction expectancies as a function of partial reinforcement; that pattern was found here, but only for the *S*s who participated in the winnings. It seems clear that if *S* did win in the situation, even though he never pulled a lever during acquisition himself, the partial effect showed clearly during extinction. It made no difference under these conditions whether *S* watched someone else play or was merely told what someone else had done while playing.

The theory that prompted this study predicted that the vicarious acquisition of the partial effect would occur, and the theory is to this extent confirmed. It was, however, originally expected that the Watch Only and the Explain Only groups would also show the partial effect. That they did not can possibly be attributed to the subtlety of the partial effect. It is an effect which is not under *S*'s verbal control, in that most *S*s do not know about the effects of partial reinforcement. One of the writers has asked several large classes to predict which of two persons would persist longer in the face of failure in a gambling situation; the one who won all the time or the one who won part of the time and irregularly. The classes split about half-and-half in their predictions. Apparently common sense does not make most people aware of what has become one of the firmest laws of psychology. It seems clear, nevertheless, that a nonverbalized experience can be imparted vicariously.

Another reason why the nonwinning groups did not show the partial effect may lie in the lower degree of participation which these groups experienced. The theory ties the partial effect closely to a mediating response, and it was guessed that this mediating response is emotional (motivational). It may well be the case that when *S* actually wins he is more of a participant in the situation and the emotional responses from winning and not winning are more likely to occur. And they must occur, it is assumed here, in order for them to get associated with the instrumental response.

Although this experiment was stimulated by a mediational response theory, and it is believed that the data are not inconsistent with such a theory, it cannot be doubted, nevertheless, that other theoretical points of view may also be able to handle the data, perhaps as well. Further, there are problems presented by these data that mediational response theory offers no clear-cut interpretation for. Other experiments are in progress in the attempt to solve some of them.

SUMMARY

Ten groups (five pairs) of subjects worked a modified slot machine, half under a 100% reinforcement schedule and half under a 25% reinforcement schedule. The first eight trials constituted the acquisition series, and the remainder of the trials were extinction. One pair of groups watched *E* play the acquisition, winning when *E* did. Another pair watched *E* play the acquisition series but did not participate in the winning. A third pair was told trial-by-trial, about a hypothetical person's acquisition series, winning when the hypothetical person won. A fourth pair was also told about the hypothetical person's play but did not share in the winning. The *S*s of a fifth and final pair played the acquisition series themselves. All *S*s then played the extinction series, being told that they could play as long as they wished.

The results showed that vicarious learning did take place as long as *S*, himself, actually participated in the winning during acquisition. The explanation of the results was offered in terms of a mediational response.

REFERENCES

1. HULL, C. L. *Principles of behavior*. New York: Appleton-Century, 1943.
2. KENDLER, H. H., KARASIK, A. D., & SCHRIER, A. M. Studies of the effect of change of drive: III. Amounts of switching produced by shifting drive from thirst to hunger and from hunger to thirst. *J. exp. Psychol.*, 1954, 47, 179–182.

3. LAZOWICK, L. M. On the nature of identification. *J. abnorm. soc. Psychol.*, 1955, **51**, 175–183.
4. LEWIS, D. J., & DUNCAN, C. P. Effect of different percentages of money reward on extinction of a lever-pulling response. *J. exp. Psychol.*, 1956, **52**, 23–27.
5. LEWIS, D. J., & DUNCAN, C. P. Expectation and resistance to extinction of a lever-pulling response as a function of percentage of reinforcement and amount of reward. *J. exp. Psychol.*, 1957, **54**, 115–120.
6. LEWIS, D. J., & Duncan, C. P. Expectation and resistance to extinction of a lever-pulling response as a function of percentage of reinforcement and number of acquisition trials. *J. exp. Psychol.*, 1958, **55**, 121–128.
7. MILLER, N. E., & DOLLARD, J. *Social learning and imitation*. New Haven: Yale Univer. Press, 1941.
8. OSGOOD, C. E. The nature and measurement of meaning. *Psychol. Bull.*, 1952, **49**, 197–237.
9. OSGOOD, C. E. Behavior theory and the social sciences. *Behav. Sci.* 1956, **1**, 167–185.
10. SPENCE, K. W., BERGMANN, G., & LIPPITT, R. A study of simple learning under irrelevant motivational-reward conditions. *J. exp. Psychol.* 1950, **40**, 539–551.
11. SPENCE, K. W. *Behavior theory and conditioning*. New Haven: Yale Univer. Press, 1956.
12. WILSON, W., WEISS, E. J., & AMSEL, A. Two tests of the Sheffield hypothesis concerning resistance to extinction, partial reinforcement, and distribution of practice. *J. exp. Psychol.*, 1955, **50**, 51–65.
13. WYCKOFF, L. B., JR. The role of observing responses in discrimination learning. *Psychol. Rev.*, 1952, **59**, 431–442.

Received August 23, 1957.

THE EXPRESSION AND INHIBITION OF HOSTILITY IN IMAGINATIVE AND OVERT BEHAVIOR[1]

JACK E. HOKANSON
University of Wisconsin
AND JESSE E. GORDON
Montana State University

SEVERAL recent studies (1, 2, 4) have used extensions of Miller's approach-avoidance paradigm to deal with behavior on projective thematic materials. In general, the findings of these studies reveal that with increased cue relevance and/or heightened motivation, there is a diminution in the projected expression of the particular motives being studied. The explanation offered by these authors is that the anxiety aroused by the heightened stimulation produces an inhibitory effect on the direct expression of the motive in the subject's (*S*'s) fantasy productions. The motives studied in this framework were hunger (2), sex (1), and threat avoidance (4).

The present study attempts to investigate the expression of hostile motivation in both imaginative and overt behavior, using the goal gradient formulation as a source of hypotheses. Three sets of variables pertaining to the expression of hostility in imaginative behavior were studied in a factorial design; in a parallel arrangement, these same variables were studied with regard to the overt expression of hostility. The variables studied were: (*a*) high and low provocation to anger (the Arousal variable), (*b*) high and low cue (a TAT card that normally elicits hostile fantasy vs. one that does not), (*c*) individual differences in the tendency to admit to or express hostility in a paper-and-pencil test situation (Expressor variable).

There are two goal-gradient models on which predictions might be based; these models lead to opposite predictions. The first alternative (Fig. 1) depicts Low Expressors as being at Point A. As the cue or motivational conditions for hostility are increased to A′, this group of *S*s should manifest a net increase in the expression of hostility; they have not yet reached a point where the inhibitory gradient is operative. This model depicts High Expressors as being at Point B. As the cue or motivational conditions for hostility are increased to Point B′, this group should show a net decrease in the expression of hostility, since the inhibitory gradient is assumed to be rising sharply in this area. Thus one would predict from this model that under conditions of high cue and arousal, people who do not ordinarily readily admit to hostility should increase their hostile expression, and those who readily admit to hostility should mainfest a decrease in hostile expression. This model assumes that people who do not readily admit to hostile feelings have less hostile motivation (are further from the goal) than those who report hostile feelings.

An alternative, though somewhat less parsimonious, model leads to opposite predictions (Fig. 2). In this formulation, the High and Low Expressors are assumed to have equal motive strengths for hostility but to differ in the degree to which they inhibit expression of this hostility. Thus both groups are considered to be at Point A, but the inhibitory gradients of the two groups differ. The inhibitory gradient of the Low Expressors is already operative, allowing for their tendency to not admit hostility readily on a paper-and-pencil test. As the cue or motivational conditions are increased to Point B, these Low Expressors should inhibit even more. The inhibitory gradient of the High Expressors, on the other hand, is not yet operative at Point A. Thus these *S*s readily admit to hostility on the paper-and-pencil test. As cue and arousal conditions are heightened, these *S*s are moved to a point where their inhibitory gradient is still low. Thus these *S*s should manifest an increase in hostile expression as they are moved to the left on the abscissa.

Exactly opposite predictions therefore follow from these models, and two-tailed tests are called for. While requiring the additional assumption that High and Low Expressors

[1] This research was conducted at the University of Wisconsin.

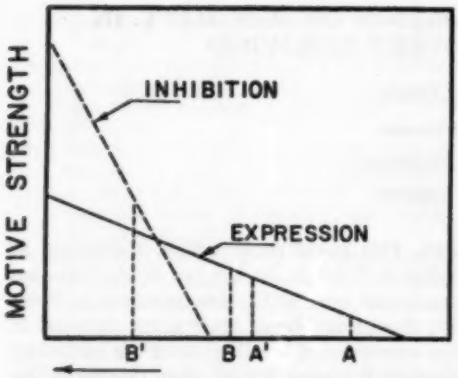

FIG. 1. MODEL BASED ON BETWEEN-GROUP DIFFERENCES IN MOTIVE STRENGTH

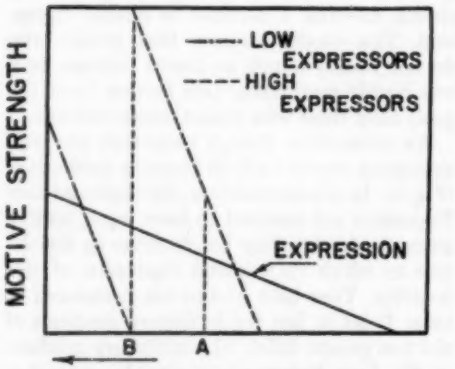

FIG. 2. MODEL BASED ON BETWEEN-GROUP DIFFERENCES IN AVOIDANCE GRADIENTS

differ in their inhibitory gradients, the second model seems to agree with the clinical observations that people who are very docile and nonaggressive are often inhibiting strong underlying hostility. Clinical experience therefore suggests that specific hypotheses be drawn from the second model. The hypotheses to be tested are:

1. People classified as Low Expressors manifest a net decrease in hostile expression, in both imaginative and overt behavior, as the cues and arousal conditions for hostility are increased.

2. People classified as High Expressors manifest a net increase in hostile expression, both in imaginative and overt behavior, as the cues and arousal conditions for hostility are increased.

The general outline of the experimental design is as follows: The extreme high and low scorers on Segal's Manifest Hostility Scale (5) were classified as High and Low Expressors. Half of each group was placed in a situation designed to arouse strong hostility and the other half placed in a low arousal situation. Hostility was measured from TAT protocols administered subsequent to arousal. Cards that usually elicit considerable hostile fantasy (High Cue) and those that do not (Low Cue) were administered to each *S*. Hostility was also measured in a behavioral situation in which *S* was given the option of administering an electric shock to *E*. Here again a device was used to provide high and low cue stimulation.

Thus a three-way classification analysis of variance was applied to both the fantasy and the overt measures of hostility. Each *S* appeared in both High and Low Cue conditions; a repeated measures analysis was used to take these correlated measures into account. The predictions made should be reflected in Expressor × Arousal and Expressor × Cue interactions.

METHOD

Forty male white undergraduate students at the University of Wisconsin taking an introductory course in psychology served as *S*s in this experiment. Their ages ranged from 18 to 26. *S*s were selected on the basis of their scores on the Segal Manifest Hostility Scale. The 20 *S*s with low scores (below 8) formed the Low Expressor group, and the 20 *S*s with high scores (above 18) comprised the High Expressor group. Segal's Manifest Hostility Scale, an objective test based on MMPI items, is described in detail elsewhere (5).

Each *S* was seen individually for approximately 40 minutes. Half of the *S*s in each group were put into an hostility-arousing situation (High Arousal), and the other half put into a neutral situation (Low Arousal). In the High Arousal condition, *S*s were asked to listen to a case study of a sadistic juvenile delinquent. Care was taken to minimize the possibility of sympathetic reactions to the delinquent, and to avoid the implication that the delinquent was mentally ill. From the *S*'s point of view, the case study portrayed a person who "simply enjoyed hurting others." This procedure was designed to produce feelings of punitiveness and hostility toward the delinquent. A case history of a juvenile delinquent who was portrayed as an innocent victim of circumstance and environmental forces was

read to the Low Arousal group. The presentation was designed to elicit sympathy for the delinquent.

Immediately after the case presentation, each *S* was asked to answer three questions designed to check whether the experimental manipulations had their desired effects. These questions required *S*s to predict the punishment received by the delinquent in the case study, to indicate what punishment *S* felt to be appropriate, and to indicate how sympathetic *S* felt toward the delinquent.

Following the above procedures, *S*s in both conditions were told that the research involving the delinquent was concluded but that *E* was interested in some other things. *S* was asked to compose imaginative stories to two TAT cards. Each *S* was presented with Cards 2 and 8 BM and asked to write a story for each one, with a five-minute time limit for each card. The order of presentation of the cards was counterbalanced with respect to the Expressor groups and the Arousal conditions. Cards 2 and 8 BM were specially selected for their difference in relevance to hostility. A normative study by Eron (3) reveals that Card 8 BM normally elicits a great deal of hostile and destructive fantasy (51.31% of responses), whereas Card 2 does not usually elicit such fantasy (6.9% of responses). Thus Card 8 BM served as the High Cue card, and 2 as Low Cue.

Scoring procedures for the TAT stories were set up prior to the experiment. Cards 2 and 8 BM had been administered to a discussion section of an intermediate psychology class, and 21 themes related to hostile expression extracted from the protocols obtained. These themes were ranked by graduate students in psychology for intensity of hostile expression. The themes were then arranged according to mean ranking and assigned to seven categories of three themes each. These categories formed the criteria for an eight-point rating scale of hostile expression on the TAT, on which a score of zero signifies a story completely devoid of hostility, and a score of seven signifies intense hostility.

A group of protocols was scored independently by *E* and two other persons. Interscorer reliability with *E* was $r = .79$ and $r = .80$. It was therefore felt that the scale scoring was reasonably reliable, and *E* proceeded to score the TAT protocols of the experimental *S*s.

The method for measuring overt hostility was one in which *S* had the option of administering electric shocks to *E*. Following the TAT administration, *E* introduced this situation with the following instructions:

> The third part of this session is again something quite different. As you are probably well aware, one of the biggest sources of error in the interpretation of psychological tests is the interpreter himself. For instance, if I were to interpret your TAT stories, such factors as being in a bad mood or feeling physically ill would certainly affect my interpretation. This is the problem I am attempting to investigate in this part of the session. Now, I am going to read over your TAT stories and make a quick interpretation to myself. Then I am going to read out loud some items from a type of personality questionnaire and try to predict how I think you would answer each item. If my prediction is wrong I want you to give me an electric shock through this apparatus. If I am correct in my prediction, then simply don't do anything and I will go on to the next item. In this way I will feel physical discomfort when my predictions are wrong; I would like to see how this affects my subsequent predictions.

The questionnaire was one developed to measure attitudes toward dating at the University of Wisconsin. The items are rather straightforward, innocuous statements which appear to keep personal involvement at a minimum. Thus to either agree or disagree with a particular item should make relatively little difference to the *S*. The nature of the items were such that regardless of *S*'s feelings about a particular item, the decision to shock or not shock was mostly determined by his readiness to express hostility overtly. It was reasoned, therefore, that the number of times *S* disagreed with *E*'s prediction (i.e., number of times *S* shocked *E*) could be used as a measure of overt hostile expression.

The questionnaire consisted of 18 items on which at least 80% of the standardization groups of University of Wisconsin male students gave the same response. Each item could be answered by Agree, Disagree, or Neutral. Items in which the agreed upon response was "neutral" were not used. Each prediction was made in a direction different from that which the normative population indicated to be the prevailing attitude. In this way, each *S* was faced with 18 predictions about himself with which he most likely disagreed. Thus he was given 18 cues to shock *E*. The items were divided into High and Low Cue items in order to parallel the design used in the fantasy measures. A High Cue item was one in which *E* predicted a response exactly opposed to the prevailing attitude expressed by the normative population, and by implication, opposed to *S*'s own attitude. A Low Cue stimulus was a prediction of "neutral," when the prevailing attitude of the norm group was either Agree or Disagree.

Each *S* received 9 High Cue and 9 Low Cue "predictions." The sequence and cue value of the items were balanced systematically between the High and Low Expressors and between the High and Low Arousal groups.

RESULTS

Tables 1, 2, and 3 reveal that the reading of the case studies was successful in arousing hostility and punitiveness. Table 1 reveals a highly significant difference between the predicted punishment for the delinquent in the High Arousal case and for the delinquent in the Low Arousal case. The mean predicted punishment score in the High Arousal condition was 3.2 (a jail sentence of approximately 16 years) while the predicted punishment in the Low Arousal was 1.4 (approximately 7 years). There was no significant difference between Low and High Expressors in their estimates of punishment, nor was there a significant Expressor × Arousal interaction. This

finding suggests that the amount of arousal was similar for both High and Low Expressor groups.

Table 2 reveals that the *S*s desired to punish the delinquent in the High Arousal case significantly more severely (approximately 27 years) than the delinquent in the Low Arousal case (approximately 3½ years in prison). There were no significant differences between Low and High Expressors in their desires to have the delinquents punished, nor was there a significant Expressor × Arousal interaction.

Table 3 indicates that *S*s expressed significantly less sympathy toward the delinquent in the High Arousal case. On an eight-point rating scale of sympathy on which a high score reflects low sympathy, the delinquent in the High Arousal case received a mean of 5.15, while the delinquent in the Low Arousal case received a mean score of 2.7. Again there were no significant differences between High and Low Expressors, and the Expressor × Arousal interaction was again nonsignificant.

These consistent results indicate that the delinquent in the High Arousal group did indeed arouse considerably more punitiveness and hostility among the *S*s than he did in the Low Arousal group.

Table 4 summarizes the analysis of variance of the hostility scores obtained on the TAT. The significant Expressor × Arousal interaction is opposite in direction to that predicted in our hypotheses. These results indicate that in a low arousal situation, the High Expressors' fantasy productions contain somewhat more hostile content than those of Low Expressors. However, when arousal is increased, the relationship is reversed; the Low Expressors' fantasy increases in hostile content while the High Expressors' fantasy decreases in hostile content. In comparing the total scores of High and Low Expressors, as a main effect, no significant difference was found.

Table 4 also indicates that the cue value of the TAT cards had a significant effect on the amount of hostile fantasy elicited. The mean score on the Low Cue card was 1.95, and 2.87 on the High Cue card. The predicted interaction of Expressor × Cue was not significant.

Table 5 summarizes the analysis of variance

TABLE 1
PUNISHMENT OF DELINQUENT PREDICTED BY SUBJECTS

Source	*df*	Sum of Squares	Mean Square	*F*
Expressor	1	2.0	2.0	—
Arousal	1	32.4	32.4	14.33*
Expressor × Arousal	1	.6	.6	—
Within	36	81.36	2.26	
Total	39	116.36		

* $p < .01$, one-tailed test.

TABLE 2
PUNISHMENT OF DELINQUENT DESIRED BY SUBJECTS

Source	*df*	Sum of Squares	Mean Square	*F*
Expressor	1	6.8	6.8	2.00
Arousal	1	228.0	228.0	67.25*
Expressor × Arousal	1	5.29	5.29	1.56
Within	36	122.1	3.39	
Total	39	362.19		

* $p < .01$, one-tailed test.

TABLE 3
EXTENT OF SYMPATHETIC FEELINGS TOWARD DELINQUENT

Source	*df*	Sum of Squares	Mean Square	*F*
Expressor	1	.9	.9	—
Arousal	1	60.0	60.0	57.69*
Expressor × Arousal	1	1.6	1.6	1.53
Within	36	37.78	1.04	
Total	39	100.28		

* $p < .01$, one-tailed test.

TABLE 4
TAT HOSTILITY SCORES

Source	*df*	Sum of Squares	Mean Square	*F*
Expressor	1	.01	.01	–
Arousal	1	.31	.31	–
Expressor × Arousal	1	18.97	18.97	5.96*
Pooled *S*s groups	36	114.7	3.18	–
Cue	1	17.11	17.11	6.88**
Cue × Expressor	1	.62	.62	–
Cue × Arousal	1	.32	.32	–
Cue × Expressor × Arousal	1	.15	.15	–
Pooled *S*s × Cue/ groups	36	71.2	1.97	
Total	79	223.39		

* $p = .05$, two-tailed test.
** $p = .02$, two-tailed test

TABLE 5
SHOCK SCORES

Source	*df*	Sum of Squares	Mean Square	*F*
Expressor	1	4.5	4.5	1.18
Arousal	1	1.0	1.0	–
Expressor × Arousal	1	3.6	3.6	–
Pooled *S*/groups	36	136.9	3.80	–
Cue	1	86.1	86.1	23.2*
Cue × Expressor	1	9.1	9.1	2.45
Cue × Arousal	1	5.5	5.5	1.48
Cue × Expressor × Arousal	1	9.2	9.2	2.47
Pooled *Ss* × Cue/groups	36	133.9	3.71	
Total	79	390.5		

* $p < .01$, two-tailed test.

of the hostility scores obtained in the situation in which *Ss* had the option of administering electric shocks to *E*. The only significant effect is that associated with the cue value of *E*'s predictions about *S*. When *E* gave a Low Cue (neutral) prediction, he was shocked on the average of 3.6 times by each *S*. When *E* gave a High Cue prediction (opposite in direction to the response *S* would probably have made), he was shocked on an average of 5.5 times by each *S*. The predicted interactions (Cue × Expressor and Arousal × Expressor) were not found to be significant. However, an interesting trend is in the same direction as the Arousal × Expressor interaction of the TAT scores. This interaction suggests that in going from Low to High Cue conditions, Low Expressors increase their frequency of shocking *E* more than do the High Expressors. The mean shock scores for Low Expressors increases from 3.5 to 6.3, while the High Expressors increase from 3.7 to 5.1.

DISCUSSION

The results depicted in Tables 4 and 5 are opposite in direction to the predictions made in Hypotheses 1 and 2. The Low Expressors increased in the amount of hostile fantasy expressed when moving from Low to High Arousal. The High Expressors decreased in the amount of hostile fantasy expressed in going from Low to High Arousal. A trend in the same direction, though not significant, was found on the shock scores in comparing Low and High Cue conditions. These results indicate that the assumptions underlying the placing of the expressive and inhibitory gradients for each group were in error. It was assumed that differences in scores on the Manifest Hostility scale reflected differences in the degree of inhibition of hostile impulses. This assumption implied that the motive strength behind hostile expression was similar for high and low scorers on the scale. Thus a model (Fig. 2) was selected that involved a single approach gradient, but different avoidance gradients for the high and low scorers on the scale. The results of this experiment suggest that the assumption of different avoidance gradients was not a valid one.

The alternative model proposed in Fig. 1 assumes that high and low scorers on the Manifest Hostility scale are operating on the same avoidance gradient, and that the differences in their scores are a function of differences in motive strength. In this model, the high scorers are placed higher on the approach gradient, and thus closer to the goal. The results of this experiment could be accounted for by this model. But the model presents another serious problem. With human *Ss*, it is reasonable to expect that factors tending to inhibit hostile expression may operate throughout the entire range of cue and arousal conditions. That is, most people have probably received sufficient punishment for expressing hostility in the past that under any given set of cue and arousal conditions, some inhibitory potential is likely to be present. Figure 1, involving linear goal gradients, does not provide for this; inhibitory potential does not accrue until some time after the expressive gradient is above the zero-strength base line. A model is needed which leads to the same predictions, but also employs the concept of inhibitory potential more centrally.

In conflict studies involving the behavior of rats that have been shocked, a single avoidance motive is usually involved (escape from shock). Such a motive may be adequately represented by a straight line. However, in human *Ss* operating under hostile motivation, it is reasonable to suppose that there are a number of avoidance motives called into play, such as fear of punishment from external sources, ambivalent feelings toward the object of aggression, fear of loss of esteem from others, fear of violation of the self concept, etc. Each of these avoidance motives may be represented

by a separate inhibitory gradient, the slopes of which may vary. When summated into a combined inhibitory potential, the resultant gradient would be curvilinear rather than rectilinear.

In such a model the resultant inhibitory gradient may extend over the entire range of arousal conditions. The model takes into consideration the criticism cited above that a rectilinear gradient does not describe the inhibitory potentials that probably exist over the entire range of approach motive strengths in humans. Further, the results of this experiment are consistent with hypotheses which might be drawn from this model. Figure 3 presents this curvilinear model applied to the conditions of this experiment.

The Low Expressors are assumed to be at Point A on the basis of their scores on the Manifest Hostility scale. Those Low Expressors who are in a High Arousal situation move to Point A′, resulting in a net increase in hostile expression. High Expressors are considered to be at Point B on the basis of their scores on the Manifest Hostility scale. When placed in a High Arousal situation, these *S*s move the same distance to the left as do the Low Expressors (there were no differences between High and Low Expressors in the extent of arousal by the case studies). They therefore move to Point B′, which results in a net decrease in hostile expression.

Although the results of this experiment are consistent with either a model involving a rectilinear avoidance gradient or one involving multiple avoidance gradients resulting in a curved gradient, we feel that this latter model is more applicable to human conflict situations. This model also does without the seemingly arbitrary assumption of Miller and Dollard that the avoidance gradient must be steeper than the approach gradient. However, it does introduce other assumptions concerning the additivity of avoidance motive strengths, which must be validated empirically. Research on animals operating under multiple avoidance drives should shed light on the validity of the hypothesized curved resultant avoidance gradient.

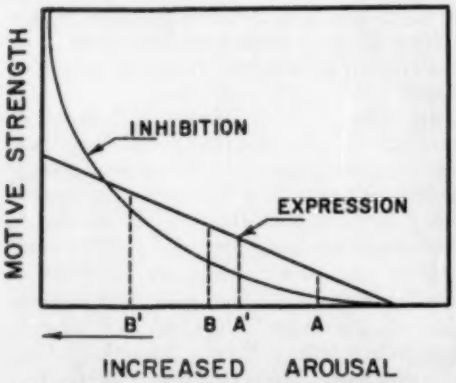

FIG. 3. MODEL BASED ON CURVED AVOIDANCE GRADIENT AND BETWEEN GROUP DIFFERENCES IN MOTIVE STRENGTH

SUMMARY

Half of a group of 40 white male college students, representing the extreme scores on the Segal Manifest Hostility Scale, were placed in a situation designed to arouse strong hostility, and half in a low arousal situation. *S*s were then permitted to express hostility in fantasy (using TAT pictures selected for differences in cue properties relevant to hostility) and in overt behavior, in a situation in which *S*s could actually hurt another person. As in the fantasy variable, the cue properties of the overt behavioral situation were varied.

Two alternative theoretical models based on Miller and Dollard's goal gradient formulation were proposed. Hypotheses derived from one of these models, which assumes differences in avoidance drive between high and low scorers on the Manifest Hostility scale, were as follows:

1. People classified as Low Expressors manifest a net decrease in hostile expression, both in fantasy and overtly, as the cues and arousal conditions for hostility are increased.

2. People who are classified as High Expressors manifest a net increase in hostile expression, both in imaginative and overt behavior, as the cues and arousal conditions for hostility are increased.

A significant Expressor × Arousal interaction in fantasy scores was in a direction opposite to that predicted, indicating that Low Expressors increased in hostile expression as arousal conditions increased, while High Expressors decreased in hostility under the same circumstances. No significant interactions involving differences in cue value were found, nor were there any significant interactions in the behavioral data.

The results of this experiment were con-

sistent with a goal gradient model in which High and Low Expressors were assumed to differ in strength of approach motivation (proximity to the goal). However, criticisms of this model were detailed, and a substitute model, involving multiple avoidance gradients which summate to produce a curvilinear gradient, was proposed.

REFERENCES

1. CLARK, R. A. The effects of sexual motivation on phantasy. In D. C. McClelland (Ed.), *Studies in motivation*. New York: Appleton-Century-Crofts, 1955. Pp. 44–57.
2. EPSTEIN, S., & SMITH, R. Thematic apperception as a measure of the hunger drive. *J. proj. Tech.*, 1956, **20**, 372.
3. ERON, L. D. A normative study of the Thematic Apperception Test. *Psychol. Monogr.*, 1950, **64**, No. 9.
4. SCOTT, W. A. The avoidance of threatening material in imaginative behavior. *J. abnorm. soc. Psychol.*, 1956, **52**, 338–346.
5. SEGAL, S. M. The relationship of hostility to authoritarianism. *J. abnorm. soc. Psychol.*, 1956, **52**, 368–372.

Received August 17, 1957.

SIZE CONSTANCY AS A FUNCTION OF PERSONAL ADJUSTMENT AND DISPOSITION[1]

NOËL JENKIN

Training School at Vineland, New Jersey

Most studies of size constancy, in relation to personality variables may be grouped into two categories. One group (2, 8, 9, 10) has claimed to show a relationship between the tendency to size constancy and some kind of personality disorder. The other group (11, 12, 15) has reported some relationship between constancy and the personality disposition of introversion-extraversion. Differing theoretical positions have instigated most of the inquiries, with the result that in each instance the independent variable has been a trait or type of diagnosis within only one or other of these two groups. There is clearly a need for further work in which variables representative both of adjustment and disposition are simultaneously studied in relation to size-distance judgment.

The direction of the relationship with a given personality variable is a matter which the literature shows to be ambiguous. Sanders and Pacht (10) found that the degree of size constancy increased concomitantly with gross degree of personality disturbance, from normality through neurosis to psychosis. Raush (9) obtained results consistent with this finding, but Lovinger (8) found that schizophrenics in poor contact with reality showed lower than normal size constancy under minimal cue conditions. Crookes (2) has also recently shown that neurotics and schizophrenics respond with a lower degree of size constancy than normals. Investigations employing the variable of introversion are somewhat more consistent, Weber (15) and Singer (11) finding that "extraverts" manifest a higher degree of size constancy than do "introverts," these results agreeing with the rather slight evidence of a similar nature earlier presented by Thouless (12).

It must be noted that there is only a broad similarity among the independent variables of each group of studies. Indeed, it may be questioned in some instances whether the variable has been correctly named. For example, Singer's (11) "thinking introverts" would be identified by Eysenck (3, 4) in terms of his "neuroticism" factor. If such a re-identification is valid, Singer's, and Sanders' and Pacht's results are in conflict, and this fact would add one more note of dissonance to an already inharmonious situation.

A possible explanation for at least some of the discrepancies lies in the nature of the instructions given to the subject (*S*). Sanders and Pacht, for example, used "objective" instructions, i.e., *S*s were "... to select the standard ... which was equal in size to the stimulus square." Singer (11) used "Look" instructions, requiring that *S* ignore field cues and avoid the use of knowledge.

The present experiment was designed with the aim of bringing partial order into this area by varying the types of instructions given *S*, distinguishing groups of poorly adjusted and normal *S*s, providing the opportunity for classifying them as introverts and extraverts, and controlling for factors of age, intelligence, and sex (12).

METHOD

Apparatus

The apparatus for obtaining measures of size-distance judgment consisted of a movable white equilateral triangle which was raised and lowered through an aperture against a black mounting board. From *S*'s viewing position, the triangle when fully lowered was obscured below the aperture. A pulley and counterweight system enabled it to be exposed in graduated degree above the aperture, as indicated by a scale at the rear. *E* sat behind the board, thus screened from *S*'s view, and manipulated the size of the triangle, which at its largest setting measured 27.5 cm from base to apex. Two white standard triangles measured respectively 8.5 cm and 6.5 cm in height. One at a time, these standard triangles were presented approximately at *S*'s eye level and directly to his left, while the movable comparison object remained directly in front of him at the same height as the standard. Lines projected upon *S* from the two stimuli would thus have made an angle of 90°. Between *S* and the comparison object was interposed a row of chairs, which reached a height well below *S*'s eye level. No objects were interposed between *S* and the standard triangle.

[1] The present investigation was designed and the data obtained when the author was at Harvard University.

Daylight was excluded, and the room was artificially lit for the experiment.

Subjects

The 40 men who served in the experiment were all students of Harvard College or of one of the Harvard graduate schools. All professed normal vision or vision corrected to normal by means of eyeglasses.[2] Ages of the *S*s ranged from 18 to 24, with a median of 20.5 years. The selection of a poorly adjusted group, numbering 20 cases, was made by an experienced clinical psychologist,[3] on the basis of referrals to the Harvard Department of Hygiene and from his own psychological examination of the cases. Various difficulties of adjustment characterized these persons and, in general, it may be said that most of them suffered from some form of psychoneurosis. The control group of 20 cases was recruited by *E* through solicitation of the services, with the offer of remuneration, of Harvard men randomly chosen. These persons were also examined by the clinical psychologist and found apparently free from symptoms of poor adjustment.

Procedure

A test battery, including Guilford's Inventory of Factors STDCR (6), the Reasoning subtest from the Chicago Tests of Primary Mental Abilities (14), and Thurstone's Concealed Figures Test (13), was individually administered to each *S* in the first of two sessions. The size-distance judgments were obtained in the second session, held on a subsequent day. The first judgment was obtained with the comparison triangle 30 ft. in front of *S* and the 8.5 cm standard triangle 15 ft. to his left. Instructions were of the "objective" type; i.e., *S* was asked to make the matches with reference to the physical size of the stimuli and to give the judgment "equal" only when he thought the two would measure the same amount. One ascending and one descending series were administered under these conditions. The standard was then placed at a distance of 2 ft. from the bridge of *S*'s nose when *S*'s head was turned directly to the left, and two more judgments were obtained. *S* and the standard triangle were then both moved to a position where *S* was 15 ft. in front of the comparison object and the standard was 12 in. to his left. After *S* had made two more judgments, the standard was moved to a position 15 ft. to his left, thus making the stimuli equidistant from him; two further judgments were then obtained.

S returned to his seat at the 30-ft. distance, and the 6.5 cm standard was placed at the 15-ft. position. *S* was now given new instructions, this time of the "analytic" type. He was asked to ignore distance and what "knowledge" he might have of the sizes. His task was to select from the distant comparison the size which would be just covered by the standard triangle if the latter was suspended in his forward line of vision and at its present distance. This definition of "retinal" equality was further illustrated by reference to the areas of space occupied by the two triangles if seen together through an hypothetical window frame at a fixed distance in front of *S*. *S* then made four more pairs of judgments at the same distances and in the same order as for the previous set of objective judgments.

RESULTS

Each major group (poorly adjusted and normal) was split at its median on the R scale of Guilford's Inventory of Factors STDCR. Since this scale is claimed to be one of the best available measures of introversion-extraversion (4, 5), it was supposed that the above procedure would isolate four subgroups which may be respectively styled "neurotic introverts," "neurotic extraverts," "normal introverts," and "normal extraverts." As it happened, the obtained range of R scores extends more towards the extreme introversion end of the scale than towards the extreme extraversion end. Strictly, therefore, the extravert groups should be given the designation "relative extravert." Table 1 shows for each subgroup the mean and *SD* on the R scale, and also indicates the means of the size matches for the various distances and the two conditions of judgment.

Preliminary analysis showed heterogeneity of variance among the measures of size-distance judgment, and they were therefore given a logarithmic transformation. Separate analyses of variance for analytic and objective

TABLE 1

EXPERIMENTAL CONDITIONS, MEAN SIZE MATCHES IN CM AND MEAN R SCORES

Instructions, standard size	Distances (ft.)		Groups			
	Comparison	Standard	Neurotic (N = 20)		Normal (N = 19)	
			Introv.	Extrav.	Introv.	Extrav.
Objective, 8.5 cm	30	15	8.45	8.25	8.40	9.08
	30	2	7.48	7.72	7.78	8.25
	15	1	7.52	8.10	8.37	8.47
	15	15	8.49	8.70	8.75	8.83
Analytic, 6.5 cm	30	15	9.93	8.17	8.08	8.47
	30	2	22.40	17.45	14.27	13.97
	15	1	23.90	18.47	18.24	14.21
	15	15	6.60	6.49	6.50	6.83
R score: Mean			19.70	32.30	26.90	43.20
SD			5.22	7.10	5.13	7.00

[2] After completion of the experiment it was found that one *S* had incorrectly made this claim and in fact suffered from uncorrected strabismus. His size-distance judgments were therefore excluded from the analysis.

[3] The author is indebted to C. C. McArthur of Harvard University for examining the cases and making the diagnoses.

conditions respectively were performed on these data and the results are shown in Table 2. It will be seen that for analytic conditions "neuroticism" is a significant source of variance, neurotics more than normals tending to match the visual angle. The *F* value for introversion-extraversion is not large enough to warrant rejecting the null hypothesis. It must be remembered, however, that the groups were not selected for introversion-extraversion as they were for neuroticism. Examination of Table 1 shows that the two traits are not entirely independent, neurotics tending to have lower R scores than normals.

Comparison of the different pairs of subgroups yields clarification. Analysis of variance of the analytic judgments by the two neurotic subgroups yielded an *F* value of 4.11 with 1 and 18 *df*, which value just misses significance at $P = .05$. The tendency indicated is for introverts to make their matches more in the direction of visual angle than in the direction of object size, and vice versa for the extraverts. Similar analysis of the two normal groups revealed no such tendency ($F = 0.08$). Only one pair of groups revealed an unequivocally significant difference—neurotic introverts and normal extraverts. The *F* value for between-group variation is 10.09 with 1 and 17 *df*, $P < .01$, and for the Distances by Groups interaction is 6.05 with 2 and 34 *df*, $P < .001$. Neurotic introverts thus make matches more in the direction of visual angle and are more responsive to changes in distance conditions. The matches of the normal extraverts tend toward or beyond the object size and tend more to remain stable despite changes in distance conditions.

It is clear, therefore, that while neuroticism appears to be the major source of between-group variation, it must be admitted as a distinct possibility that introversion is a contributing factor in analytic size-distance judgments.

Analysis of variance of the objective judgments, also summarized in Table 2, shows no significant differences between groups and no significant interactions. The tendency present, under these conditions, is opposite to that for analytic judgments, normal persons making responses slightly more in the direction of visual angle than the neurotics. The latter, at the disparate distances, uniformly showed "overconstancy."

The results described above are obtained with adequate control of factors previously shown to be possible functions of size-distance judgment, namely, sex and age (12). A further possibility is that the significant difference in analytic size-distance judgments between neurotic and normal groups may be accounted for in terms of intelligence (12). Analysis of results from the two measures of intellectual ability yields no reason to believe that this is the case (Reasoning, $t = 1.02$, $P > .05$; Concealed Figures, $t = 0.85$, $P > .05$).

Data from the equidistant judgments provide a control on the possibility that the difference discovered is not due to distance but rather a constant error of size estimation such as found in other work on affective processes in perception (7). Inspection of Table 1 shows that there was an over-all tendency for all *S*s slightly to overestimate the size of the standard object at the 15-ft./15-ft. position. Analysis of variance of the data showed, however, no significant differences between the groups and no significant interactions.

TABLE 2

ANALYSES OF VARIANCE OF SIZE MATCHES

	df	Analytic		Objective	
		Mean Square	*F*	Mean Square	*F*
Neuroticism (N)	1	.4563	6.32**	.0152	2.49
Introversion (I)	1	.1658	2.30	.0085	1.39
N × I	1	.0529	0.73	.0004	0.07
Subjects[a]	35	.0722		.0061	
Distances (D)	2	.9601	72.73****	.0161	7.67***
N × D	2	.0573	4.34*	.0007	0.33
I × D	2	.0223	1.69	.0001	0.05
N × I × D	2	.0061	0.46	.0039	1.86
Residual	70	.0132		.0021	

* $P < .03$.
** $P < .02$.
*** $P < .01$.
**** $P < .001$.

[a] Since one *S* was lost from the normal extravert group (see text, footnote 3) his scores were replaced by the appropriate cell means. One *df* was then subtracted from the "Subjects" classification and 3 *df* were subtracted from the total.

DISCUSSION

The results of this experiment clarify, to some extent, the ambiguity resulting from studies by previous investigators. In the present experiment, poorly adjusted people behaved like Singer's "thinking introverts"

(11). As previously mentioned, Eysenck's factorial work (3, 4) suggests that Singer's so-called introverts actually are better described with reference to the neurotic dimension of personality. In the present experiment, the tendency of poorly adjusted persons to judge in terms of the visual angle was true only for analytic intructions, similar in intent to the "Look" instructions of Singer. Under objective instructions, these people showed an opposite tendency, i.e., toward constancy or "over-constancy," relative to the performance of normal persons. While this unexpected result does not reach statistical significance, it is interesting to note that it is in a direction consistent with the finding of Sanders and Pacht (10) who also used objective instructions. It is clear that no generalized statement can be made about the "tendency to size constancy" of psychoneurotic or poorly adjusted persons, without specifying the nature of the instruction-produced task attitude. Previous studies, together with the present results, suggest that the tendency is smaller than that of normal persons in the analytic situation and greater than that of normals in the objective situation.

The complexity of the problem is further realized when the role of introversion is considered. The present results make it difficult to resist the conclusion that neuroticism and introversion appear to be additive in their effect on analytic size judgment. Introverted neurotics differ markedly from extraverted normals in their analytic judgments, whereas normal introverts differ from neurotic extraverts only to an insignificant degree.

Singer (11) has adopted a position similar to that of Weber (15) in suggesting that the introvert, because of immersion in self and analysis of his own cognitive processes, is less dependent than the extravert on the stimulus characteristics of the external environment and therefore shows less tendency toward constancy. The present type of analysis shows that such a view is too simple. Introversion by itself was in the present experiment insufficient to produce the effect. Singer's reasoning is still plausible if it is made to refer to introverts who are also neurotic or, in other words, to Eysenck's "dysthymic" category. However, his explanation cannot account for the opposite tendency under objective conditions of judgment, found as a nonsignificant trend in this experiment, and found to a significant degree by Sanders and Pacht (10).

While there is danger in extrapolating from the present results to a different type of population, such a venture may be attempted in view of evidence and theory already existing in the literature. If we admit that the disease of schizophrenia represents an extreme both of introverted disposition and of maladjustment, it may be supposed that the victims of this malady would surpass even the introverted neurotics of the present study in their analytic tendency to match the visual angle. This suggestion may do violence to the widely held view of psychosis as a factor independent of neurosis, but it is consistent with the finding of Crookes (2) that neurotics have lower constancy indices than normals and that schizophrenics show indices significantly lower than either of the other groups. Also pertinent is Lovinger's (8) finding, cited earlier. In the meantime, therefore, the best theoretical rationale for the existing evidence seems to be an extension of Bruner's (1) suggestion that schizophrenia represents a withdrawal from object relations and therefore involves a breakdown in the constancies. Such a withdrawal from object relations may also occur to a lesser degree in neurotic introverts. The results inconsistent with such a theory may at present be attributed either to presentation of the size-judging situation as an objective rather than an analytic task (e.g., 10) or to a failure adequately to communicate the analytic nature of the task. Further study of the objective task is needed to yield understanding of why it has properties different from the analytic task in relation to adjustment and disposition.

SUMMARY

A size-distance judging task was given to 40 men who were classified in four groups, respectively styled "neurotic introverts," "neurotic extraverts," "normal introverts," and "normal extraverts." Analysis of data from four distances under two conditions of judgment, i.e., objective and analytic, indicated that neuroticism was the major source of between-group variation. Under analytic conditions, neurotic persons tended to match the stimulus in terms of visual angle, and normals in terms of object size. The data also

indicated that introversion to a less marked degree was also related to matches in the direction of visual angle. Objective judgments showed the opposite trend to a nonsignificant degree. The results were discussed in relation to previous research in which the several variables of the present design were separately treated with some inconsistent outcomes.

REFERENCES

1. BRUNER, J. S. Personality dynamics and the process of perceiving. In R. R. Blake & G. V. Ramsey (Eds.), *Perception: An approach to personality*. New York: Ronald, 1951. Pp. 121–147.
2. CROOKES, T. G. Size constancy and literalness in the Rorschach Test. *Brit. J. Med. Psychol.*, 1957, **30**, 99–106.
3. EYSENCK, H. J. *The scientific study of personality*. New York: Macmillan, 1952.
4. EYSENCK, H. J. *The structure of human personality*. New York: Wiley, 1953.
5. FRANKS, C. M. Conditioning and personality: A study of normal and neurotic subjects. *J. abnorm. soc. Psychol.*, 1956, **52**, 143–150.
6. GUILFORD, J. P. *Inventory of factors STDCR*. Beverly Hills, Calif.: Sheridan Supply, 1940.
7. JENKIN, N. Affective processes in perception. *Psychol. Bull.*, 1957, **54**, 100–127.
8. LOVINGER, E. Perceptual contact with reality in schizophrenia. *J. abnorm. soc. Psychol.*, 1956, **52**, 87–91.
9. RAUSH, H. L. Perceptual constancy in schizophrenia. I. Size constancy. *J. Pers.*, 1952, **21**, 176–187.
10. SANDERS, R., & PACHT, A. R. Perceptual size constancy of known clinical groups. *J. consult. Psychol.*, 1952, **16**, 440–444.
11. SINGER, J. L. Personal and environmental determinants of perception in a size-constancy experiment. *J. exp. Psychol.*, 1952, **43**, 420–427.
12. THOULESS, R. H. Individual differences in phenomenal regression. *Brit. J. Psychol.*, 1932, **22**, 216–241.
13. THURSTONE, L. L. *A factorial study of perception*. Chicago: Univer. Chicago Press, 1944.
14. THURSTONE, L. L., & THURSTONE, THELMA G. *Chicago Tests of Primary Mental Abilities*. Chicago: Science Research Associates, 1947.
15. WEBER, C. The relation of personality trends to degrees of visual constancy correction for size and form. *J. appl. Psychol.*, 1939, **23**, 703–708.

Received August 23, 1957.

VALUE DIFFERENCES AMONG SOUTH AFRICAN STUDENTS

KURT DANZIGER

University of Natal

ABERLE (1) has indicated that while "contemporary social theory . . . strongly emphasizes the crucial importance of a shared system of ultimate values as an element in any society" it tends to gloss over the diversity of subsystems that exist within any complex society. The South African situation is of interest in this connection as it provides an opportunity for studying value conflict under conditions of marked social tension.

The white minority in South Africa exercises a virtual monopoly of political, economic, and military power, and most of its elected representatives are committed to a policy of maintaining white supremacy at any cost. On the other hand, as the nonwhite majority is drawn increasingly into the web of a modern industrial civilization, its reaction to this state of affairs tends to assume a more and more determined character. The details of this conflict do not concern us here and may be gathered from a number of sources, e.g., Kuper (7, 8, 9).

Among the nonwhite section of the South African population the great majority are Bantu Africans, but there are also people of mixed descent and an Indian community which includes the most important nonwhite middle-class element. The white population is divided into English-speaking and Afrikaans-speaking sections.

In a homogeneous community the members share a common definition of their social situation, to use W. I. Thomas' term. They agree about their position in the world, their goals and how to achieve them; they have a similar evaluation of their society as a whole and of their position in it. In a society split by internal conflict this is not the case. The same social situation is defined very differently by the members of the opposing groups. What is regarded as a just and desirable social order by some is detested by others. The less agreement there is about the characteristics of the social situation in which the different groups find themselves the more intense the underlying conflict is likely to be. The first task of a study of value differences in a situation of social conflict is therefore the measurement of group differences in the definition of the social situation. In the South African context this means examining the evaluation that whites and nonwhites make of the dominant pattern of their society as it exists at present.

But the subjective manifestations of social conflict are not exhausted by differences in the definition of the social situation. They enter much more deeply into the personality and serve to determine the inner values of the individual in important ways. When people differ markedly in their conception of the society in which they live, they are also likely to differ significantly in the goals which they set themselves within that context. The individual who feels himself to be a member of an unjustly oppressed group is likely to have different personal aims and values than the individual who denies the existence of social oppression in his society. Differences in the definition of the external social situation thus involve concomitant differences about the role and duty of the individual in this situation. The second part of the present investigation is therefore concerned with assessing some of the main differences that arise between white and nonwhite South Africans in this respect.

METHOD

Subjects

The *S*s of the present investigation were 195 students at a number of higher educational institutions in South Africa. The total number was made up of 81 white university students, 34 white students at a technical college, 37 Indian and 43 African university students. The technical college students were only involved in the first part of the investigation. Of the white *S*s, 21 were Afrikaans-speaking. The socioeconomic level of the white group as a whole was considerably above that of the nonwhite group. The university students included in the sample were all attending introductory psychology or sociology courses in segregated classes and proceeding to arts, sci-

ence or medical degrees. The technical college students were attending a course in pharmacy.

Questionnaires were administered on a voluntary basis in the classes attended by the *S*s. There was no consultation among individual *S*s.

Characterization of "White Civilization"

In the first part of the investigation, an attempt was made to assess the *S*s' over-all reaction to the social situation in South Africa by measuring their reaction to the key term "White Civilization." This term is frequently used in political and other discussions, both public and private, also in speeches and publications at all levels. Generally the context is one of a call for the preservation of "White Civilization." The term in fact refers to the ruling pattern of social relationships in the country. Thus, to ask people to characterize this entity is to ask them to define the dominant social pattern in their society. In this way it is possible to get an assessment of a key element in their definition of the social situation in which they live. In a different national context, the same purpose might have been served by getting characterizations of the meaning of such terms as "democracy," "free enterprise," "the American Way of Life," etc.

*S*s were presented with a list of 14 features which, according to the instructions, "different people have claimed to be highly characteristic of white civilization in South Africa." They were asked to indicate which of these features they considered to be really characteristic of white civilization and which not. They were also asked to rank the former in order of importance. The items available for choice were:

1. High standards of morality in the sphere of family life
2. Respect for law and order
3. Unjust oppression of nonwhite people
4. Great scientific achievements
5. A superior system for the production of material goods
6. Exploitation of cheap black labor
7. Arbitrary and undemocratic government
8. An unfair distribution of material goods
9. An emphasis on the rights and liberties of the individual
10. The ideal of equality of opportunity for all
11. Brutality towards the underdog
12. Hypocrisy in dealing with nonwhites
13. A legal system based on justice and fairplay
14. The ideal of peace among all men

Personal Values

On the hypothesis that differing evaluations of the social situation would be associated with preferences in respect to individual goals, *S*s were asked to answer the following open ended questions:

1. What two things would you like to have that you don't now have?
2. What two things would you most like your child to have?
3. What two things could you conceivably accomplish during your lifetime that you would be most proud of?
4. For what end would you be willing to make the greatest sacrifice of personal confort, time, and money?
5. What are the two worst things that could happen to you during your lifetime?
6. As a parent, what two specific lessons will you try hardest to teach your children?

The goals mentioned in reply to the first four questions were categorized as follows:

Political goals: e.g., "a new system of government," "the right to vote," "freedom of speech and movement," "the right to choose how I am governed," "political freedom," "liberation of my people," etc.

Educational goals: in the sense of a desire for improved educational opportunities, not of personal success

Specific helpfulness: e.g., "help my people," "give medical service to poor Africans," "establish a clinic in an African area," etc.

Abstract helpfulness: e.g., "reducing human unhappiness," "help those in need," "creating peace and universal goodwill," etc.

Familial goals: e.g., "a happy family life", "raising a family" etc.

Material possessions: e.g., "money," "a house," "a motor car," etc.

Personal characteristics, exploits and achievements: e.g., "ambition," "shining at sports," "getting my degree," "health," etc.

The categories used in analyzing responses to the last two questions were somewhat different and are discussed below.

Abbreviated F Scale

An abbreviated form of the F scale was administered to the white university students only. This consisted of Items Nos. 1, 4, 8, 9, 18, 19, 26, 27, 29, 31, 34, 37, 42 selected from the list of items (Forms 45 and 40) presented by Adorno et al. (2, pp. 255–257). Each item was scored on a five-point scale, a score of 0 representing strong disagreement and a score of 4 representing strong agreement.

RESULTS AND DISCUSSION

Characterization of "White Civilization"

The 14 choice items presented as possible characteristics of "White Civilization" actually include eight items (Nos. 1, 2, 4, 5, 9, 10, 13, 14) that refer to favorable characteristics and six items (Nos. 3, 6, 7, 8, 11, 12) that refer to unfavorable characteristics. The proportion of favorable to unfavorable items in the choice of each *S* may be taken as an over-all measure of his reaction to the term "White Civilization." In assessing this proportion, only the first five choices for each *S* were taken into account. By means of this procedure, *S*s may be readily divided into three groups: (*a*) Those who choose not more than one negative item in their first five choices—these *S*s are clearly favorably disposed to the social situation designated by the stimulus term; (*b*) those *S*s who choose not more than one positive item in their first five choices and whose reaction is therefore unfavorable on the whole; (*c*) those *S*s who choose more than one positive *and* more than one negative item and who may be regarded as being intermediate in their reaction to the stimulus term. The frequency of the three response types for each of the experimental groups is shown in Table 1.

TABLE 1
FREQUENCY OF DIFFERENT OVER-ALL RESPONSES TO "WHITE CIVILIZATION"

Response Category	White[a]		Nonwhite[a]	
	University Students	Technical College	Indians	Africans
Favorable	22	23	5	3
Intermediate	27	6	9	5
Unfavorable	23	5	23	33
Total	72	34	37	41

[a] Nine white university students and two Africans returned the questionnaire unanswered.

The table shows (*a*) that whites tend to evaluate the stimulus more favorably than nonwhites; (*b*) that the proportion of favorable evaluations is much less among white university students than among technical college students; and (*c*) that the proportion of unfavorable evaluations may be slightly greater among African than among Indian *S*s. These differences can be related to marked differences in the social situation of the different groups. The white *S*s as a whole tend to evaluate favorably a social order of which they are the beneficiaries, while the nonwhite *S*s evaluate negatively an order that reduces them to the level of second-class citizens. On the one hand, the situation is defined primarily in terms of certain positive characteristics; on the other hand, the situation is seen as threatening and repressive.

The difference between white university and technical college students and the unusually high proportion of "intermediate" response types among the former may be a reflection of a conflict that exists for the university students between the stereotyped attitude of their homes and a more critical attitude to which they are exposed at the university. In this and other respects the technical college students are much more representative of the white population as a whole; an analysis of white voting behavior certainly indicates an overwhelmingly positive evaluation of "White Civilization."

The difference between African and Indian students may be a function of the objective fact that under the existing social order the former have even fewer civil rights than the latter.

Nature of Favorable Characteristics Chosen

Further light is thrown on the over-all difference between white and nonwhite *S*s by an analysis of the *nature* of the favorable items that are chosen by each group as being characteristic of "White Civilization." The most popular items for both white groups are Nos. 1, 2, and 13—"high standards of morality in the sphere of family life," "respect for law and order," and "a legal system based on justice and fairplay." It is noteworthy that all three items, and particularly the first and the last, would appear to establish some claim as to the *moral* advantage of "White Civilization." On the other hand, insofar as nonwhite *S*s choose any favorable items at all, these tend to be items concerned with the *material* advantage of "White Civilization," namely, Items 4 and 5, "great scientific achievements" and "a superior system for the production of material goods." The difference between the groups in this respect is presented in Table 2.

Thus, while the order of which they are the privileged representatives appears to the white *S*s as moral and just, the nonwhite *S*s are prepared to concede at most material advantages

TABLE 2

NUMBER OF CHOICES OF FAVORABLE "MORAL" AND "MATERIAL" CHARACTERISTICS OF "WHITE CIVILIZATION"

	White	Non-white
Choice of "moral" characteristics, Items 1 and 13.	67	9
Choice of "material" characteristics, Items 4 and 5.	15	21

Note.—$\chi^2 = 27.5$; $p < .01$.

TABLE 3

RESPONSES TO THE QUESTION: "WHAT TWO THINGS WOULD YOU LIKE TO HAVE THAT YOU DON'T NOW HAVE?"

Response Category	White	Non-white
Political goals	Nil	61
Educational opportunities	4	15
Familial goals	33	5
Personal qualities and achievements	69	16
Total	106	97

to the same order, rejecting entirely its claims to moral excellence. Thus the latter group defines the dominant social pattern as basically immoral and unjust. The items it most frequently chooses as characteristic of "White Civilization" are Nos. 3, 6, and 7, "unjust oppression of nonwhite people," "exploitation of cheap black labor" and "arbitrary and undemocratic government."

Certain differences of attitude within the white group may be mentioned briefly. While the English-speaking *Ss* often accept at least one unfavorable item as being characteristic and reject a few favorable items as not being characteristic of "White Civilization," the Afrikaans-speaking *Ss* show an extremely rigid tendency always to accept only favorable characteristics and to reject only unfavorable characteristics. For these *Ss* therefore, the term "White Civilization" acts as a true stereotype towards which no conscious ambivalence can be tolerated.

A difference that arises between English-speaking university and technical college students relates to the area in which negative characteristics are admitted. Both English-speaking groups tend to accept Item 6—"exploitation of cheap black labor," and to reject Item 10—"ideal of equality of opportunity for all." But only the university students, to any extent, accept Item 7—"arbitrary and undemocratic government"—as characteristic, and reject Item 9—"an emphasis on the rights and liberties of the individual" as not characteristic of "White Civilization." This finding suggests that while most English-speaking *Ss* are prepared to concede the existence of a certain amount of unfair discrimination in the economic sphere, only the university students are prepared to extend criticism to the political sphere. Insofar as they do this they are presumably applying democratic standards that are foreign to the ordinary white citizen.

Personal Values

The second part of the investigation was concerned with differences of personal values between white and nonwhite groups. For this part of the study, the technical college students were not available, and the results are therefore based on university students only.

Replies to the open-ended questions were scored in accordance with the categories described earlier. The frequency of all the response categories among white and nonwhite *Ss* was tabulated separately for each of the six questions. In each case a value of χ^2 was obtained that was significant at the .01 level. Large and consistent differences were found to exist between the two groups. Table 3 shows the main differences that occurred in the responses to the first question.

The difference between the two groups is striking. Whereas the white group concentrates its choices in the sphere of purely personal satisfactions, the nonwhite group tends to put wider social interests first. The same pattern of group differences occurs with the responses to each of the six questions, and in each case, the difference is highly significant statistically. The distinction between the response pattern of white and nonwhite groups appears to be similar to one made by Allport and Gillespie (4) who use the term *privatism* to refer to the American and often the European student's concern with "private values" and his absorption in the "goals of private worldly success." By contrast, Egyptians, Mexicans, Japanese, and South African Bantus were found to be far more orientated towards political and social issues. Similarly, Hughes and Thompson (6) in the United States found that Northern

white children put a lower value-emphasis on justice and group unity than Southern Negro children. Lee (10) and others have stressed the basic nature of individualistic values in the culture patterns of societies generally designated as "Western."

One reason for this difference probably lies in the fact that certain societies have a strong tradition of economic individualism, whereas others do not. But the sociopolitical orientation of certain groups is not simply negatively conditioned by an absence of individualistic values; it is a result of certain positive pressures towards redressing real and perceived limitations on the group by means of group action. Members of national and class groups whose aspirations and "path of free movement" are limited simply because of this group affiliation tend to internalize the social aspirations of the group so as to turn them into individual aspirations for each member. In the case of less privileged groups, social and individual goals tend to coincide, whereas in dominant social groups, conflict between public duty and individual interest tends to occur. While most of the nonwhite *S*s who evaluated the existing dominant social order negatively also mentioned a change in this order as one of their personal desires, not one of the white *S*s did so, although 23 of them had apparently recognized the injustice of this order. For these *S*s there exists a characteristic discrepancy between the definition of the social situation and the personal aspirations of the individual.

Further light is thrown on this pattern by comparing the frequency of the two response categories—"abstract helpfulness" and "specific helpfulness"—for Questions 3 and 4. Table 4 combines responses in these two categories for both questions. It is in accordance with the whites' basically individualistic orientation that when they do declare themselves to be willing to make sacrifices for an end other than their private satisfaction, this end tends to assume a conveniently abstract and nebulous character.

On the whole, the person with a specific social aim, like building a clinic, for example, is much more likely to translate his aim into social action than the person who aims merely to "reduce human unhappiness." The point is that social action must always be concrete and the person with an abstract aim must first give his aim a concrete content in terms of the situation that actually faces him. As long as the aim remains abstract and formal, its function may not really be that of reorientating the individual towards social action, but rather that of assuaging the guilt that arises from the conflict between social ideals and private interest. For the socially oriented person, on the other hand, social aims naturally assume a concrete content, as they arise directly out of the demands of a specific external situation that have become identified with his individual interest. Thus the nonwhite *S*s repeatedly mentioned the goal of helping their own communities also in other than political ways. In the white *S*s, the split between private and public goals may be accompanied by guilt feelings where the individual is aware of the demands of the social situation and of his own failure to live up to them. Allport and Kramer's (3) finding that their less prejudiced *S*s showed guilt about what prejudice they did possess may be relevant here.

TABLE 4
FREQUENCY OF "HELPFULNESS" RESPONSES
(Questions 3 and 4)

Response Category	White	Nonwhite
Abstract helpfulness	29	15
Specific helpfulness	6	30

Note.—$\chi^2 = 20.6$; $p < .01$.

The pursuit of private satisfactions, however, is not only the result of the expectation of positive gains in that direction; to an extent which can never be accurately determined, it is also conditioned by fear of the consequences to the individual of participation in sociopolitical action. This factor varies with the level of social repression that exists in the society. In South Africa, for example, the individual faces serious penalties should he advocate a fundamental change in the existing social pattern. The socially oriented individual is therefore compelled to face the very real possibility of political persecution with all its attendant personal consequences for him. The privatistic individual escapes this, but as Table 5 shows, he may then develop less rational fears that are characteristic for his social isolation.

Before considering the important group differences shown in Table 5, it is worthwhile noting that physical misfortunes and ill health are mentioned with the same frequency by

TABLE 5
FREQUENCY OF VARIOUS RESPONSES TO THE QUESTION: "WHAT ARE THE TWO WORST THINGS THAT COULD HAPPEN TO YOU DURING YOUR LIFETIME?"

Response Category	White	Nonwhite
Political persecution	8	32
Economic insecurity	18	31
Physical misfortune, ill health	41	38
Loss of moral standing	9	18
Death of a relative	14	9
Mental instability	13	3
Loss or lack of love	36	5
Total	139	136

both groups, thus indicating that this response is not affected by the differential social forces acting on the two groups.

The main pattern of group differences follows the differences of orientation that have already been discussed. Members of the underprivileged nonwhite group, oriented as they are towards a change in the social situation, have to face political persecution and economic insecurity as real possibilities. But this eventuality is not only unpleasant in itself, it also threatens to undermine the dominant orientation of the individual towards social goals because of the possible consequences of such an orientation. Thus a number of the nonwhite *S*s also express a fear of losing their moral standing in their community by a failure to live up to their social ideals.

On the other hand, the white *S*s often mention certain unpleasant possibilities that are almost nonexistent in the nonwhite sample. Here we get such responses as "insanity," "losing one's faith in everybody and everything," "losing all one's friends," "being unable to love anybody," "being unhappily married," "not getting married," "loneliness," etc. These responses have been classified in the last two categories of Table 5. They all appear to indicate a basic preoccupation with social isolation. The privatistic individual pays the penalty of his withdrawal into a world of individual goals in the form of a basic anxiety regarding social isolation or "loss of love." As the social relationships of his private life are not based on common social aspirations but on the capacity of each partner to furnish the other with narcissistic satisfactions, they are subject to a basic instability. His real isolation would appear to make the privatistic individual hyperdependent on constant proofs of love from his surroundings, while at the same time his basically narcissistic orientation prevents him from forming relationships of a type that would ensure his real social integration.

Insofar as the privatistic individual's relationships to other people are not founded on common goals, there will be a certain tendency to stress the more external features of these relationships and to develop a social "persona" that is suitable for such superficial relationships. Some evidence in this direction is contained in the replies to the question: "As a parent, what two specific lessons would you try hardest to teach your children?" There is a certain tendency for the white students to more frequently state that they would teach such virtues as "good manners," "politeness," "sociability," "good behavior," "the ability to get on with others," as well as "sportsmanship," "cheerfulness," and the ability "not to show one's feelings." These responses occur 36 times in the white group and 14 times in the nonwhite group. This difference is at least suggestive of a somewhat greater concern with the *external appearance* of social integration in the white group. Presumably such concern would be related to the very real difficulties which the intensely privatistic individual faces in establishing genuine social ties.

Another type of response to the last question which is of special interest is that related to authoritariansim. The main responses here were "obedience," "discipline," and "respect." They occurred 15 times in the white group and 20 times in the nonwhite group. However, it would be a mistake to assign the same significance to these responses in the two groups. In the white group these responses occurred mostly among *S*s who had a favorable over-all response to "White Civilization" (difference significant at .05 level), but no such relationship exists in the nonwhite group. Moreover, when the F scale was administered to the white *S*s, it was found that those whose over-all response to "White Civilization" was favorable tended to get higher scores on the F scale (difference significant at .02 level). The product-moment correlation between *F* score and favorability of response to "White Civilization" (scored on a six-point scale) was .65. Thus in the white group, acceptance of exist-

TABLE 6
RELATIONSHIP BETWEEN F SCORE AND RESPONSES TO "WHITE CIVILIZATION"

Response to "White Civilization"	F Scale	
	Mean per item	*SD*
Favorable	2.29	.55
Intermediate	1.84	.58
Unfavorable	1.59	.63

ing social discrimination and ethnocentrism in some cases form a pattern with an emphasis on authoritarian values in the training of children and "fascism" as measured by the F scale that is reminiscent of the "authoritarian personality" (2).

But in the nonwhite group such authoritarian values are frequently associated with an intensely critical attitude to the existing social order. They are quite likely to arise from a rational concern with the necessity for group solidarity in the face of external social threats to its aspirations. When these *S*s were asked to explain the reasons for their social orientation, references to the necessity for group solidarity in the face of social oppression were common (5). The interpretation of the pattern of "authoritarianism" must always take into account the wider social context. In the present case, the position is further complicated by the survival of traditional patterns in the nonwhite group to an unknown extent. Some estimate of the influence of these patterns could be made by examining the differences that exist between Indian and African *S*s. This has not been done in the present study which is concerned with the over-all differences between whites and nonwhites. The differences within the nonwhite group were actually very much less marked than these over-all differences. To some extent the patterns of privatism and social orientation do arise out of the traditions of the groups concerned. But they are also conditioned by the objective social situation in which the groups find themselves at the present time and by their subjective definitions of that situation.

The relationship between the real situation and its subjective definition has a different significance for the privileged white and the unprivileged nonwhite group. Insofar as the latter recognizes the elements of oppression in the objective situation, and the great majority of the *S*s appeared to do so, it arrives at a definition of the situation which makes a fusion of individual and social interest possible.

On the other hand, members of the privileged group have two alternatives. They may identify self-interest with justice and approve of the existing social situation. As that situation is in fact based on naked power and the rule of the strong, their approval of it makes them predisposed to accept an openly Fascist ideology, which is what a number of our white *S*s with high F-scale scores have actually done. On the other hand, those white *S*s who possess standards of social life that render it impossible to identify self-interest with justice are affected by an internal conflict in which self-interest always wins, although at a certain cost. Of course it is possible for this conflict to be resolved by an identification with the oppressed, but no clear case of this occurred in our sample.

South African society is marked by an exceptionally severe conflict between the process of social integration fostered by a growing modern industrial system and the process of deepening caste divisions which is initiated by the power of a state that represents the dominant caste. This conflict leads to wide divergencies of outlook between the members of the privileged and nonprivileged castes. These divergencies of outlook make communication across caste barriers exceptionally difficult, as they appear to affect not merely "surface" opinions but also more basic aspects of personality organization. Thus disagreements about social values and individual goals exacerbate the basic conflict. For example, opportunities for limited personal advancement within the existing social framework are sometimes offered by official policy in accordance with its scheme of values only to be frequently rejected with indignation by the supposed beneficiaries whose own aspirations lie in quite another direction. The result is a further deepening of tensions, except in so far as counteracting influences are brought to bear on the system from outside.

SUMMARY

Groups of white and nonwhite South African students were required to choose features that they considered to be characteristic of "White Civilization" in South Africa. Great differences in their definition of the social situation in the country were indicated by the

fact that the nonwhite students attributed mainly negative features to the stimulus term while the white students responded to it much more favorably. The latter claimed certain moral advantages for "White Civilization," while the former regarded it mainly as a system of exploitation and oppression with at best certain material advantages.

This difference in attitude to the dominant pattern of social relationships in the country appeared to be linked to a difference of personal values. In response to a number of questions designed to elicit value choices, the nonwhite students put a great deal of emphasis on political values and also on the goal of helping their own communities in other ways. On the other hand, the white students were almost entirely limited to choices in the sphere of private satisfactions—they desired success in their career, in their family life, in the cultivation of certain personal qualities, and personal exploits. Where they did pay due to the value of general helpfulness this was done in a characteristically abstract manner that contrasted with the more concrete approach of the nonwhite subjects to social improvements.

These differences may be interpreted in terms of a basic value of privatism that determines the life organization of the white subjects and contrasts strongly with the social orientation of the nonwhite subjects. In the case of the latter, there is a certain identity of social and individual aspirations, whereas the privatistic individual is marked by a split between private and public goals with the former having much greater force for him.

The privatistic and the socially oriented individuals are subject to different forms of anxiety. Whereas the latter tends to fear political persecution and a possible loss of moral standing, the latter's great fear is of social isolation in its various forms. In some cases this seems to go with an emphasis on the external appearance of social relationships instead of on their genuine content. It may be suggested that the privatistic individual has a basically narcissistic orientation which stands in the way of his forming genuine social ties based on shared aspirations. At the same time he requires proofs of love from his environment. His basically exploitative attitude to people leads to the prevalence of fears of social isolation.

One way of dealing with this situation appears to consist in the pattern of "authoritarianism" where the individual identifies his personal interest with that of the most powerful group, and submission to the powerful becomes a basic value. There was evidence of the existence of this pattern in some of our white subjects.

REFERENCES

1. ABERLE, D. Shared values in complex societies. *Am. sociol. Rev.*, 1950, **15**, 495–502.
2. ADORNO, T. W., FRENKEL-BRUNSWIK, ELSE, LEVINSON, D. J., & SANFORD, R. N. *The authoritarian personality*. New York: Harper, 1950.
3. ALLPORT, G. W., & KRAMER, B. M. Some roots of prejudice. *J. Psychol.*, 1946, **22**, 9–39.
4. ALLPORT, G. W., & GILLESPIE, J. M. *Youth's outlook on the future*. New York: Doubleday, 1955.
5. DANZIGER, K. Self-interpretations of group differences in values. *J. soc. Psychol.*, in press.
6. HUGHES, J. H., & THOMPSON, G. G. A comparison of the value systems of southern Negro and northern white youth. *J. educ. Psychol.*, 1954, **45**, 300–309.
7. KUPER, L. South African native: Caste, proletariat or race? *Soc. Forces*, 1949, **28**, 146–153.
8. KUPER, L. The control of social change: A South African experiment. *Soc. Forces*, 1954, **33**, 19–29.
9. KUPER, L. *Passive resistance in South Africa*. London: Jonathan Cape, 1956.
10. LEE, D. Are basic needs ultimate? *J. abnorm. soc. Psychol.*, 1948, **43**, 391–395.

Received August 23, 1957.

FAMILY PATTERNS CORRELATED WITH ADOLESCENT PERSONALITY STRUCTURE

ROBERT F. PECK

University of Texas

IN PRESENT day theories of personality development it is almost axiomatic that the emotional dynamics of the child's family play a dominant part in shaping his personality. Even today, however, the most detailed and persuasive evidence for this proposition comes largely from individual case studies and, chiefly, from clinical studies of maladjusted individuals, who most often seem to present a compelling reason for investigating their histories. There have been, of course, a sizable number of studies that relate certain overt traits of child behavior to similar traits in parents (1). Perhaps the most comprehensive effort to relate child behavior patterns to parental practices has been the Yale cross-cultural survey (2).

For the most part, these studies have adhered to overt behavorial characteristics that are readily observable in a common-sense way, both in children and in parents. Conceptual constructs about "inner" personality dimensions, on the other hand, have rarely been tested in any quantified way, or on more than a few individual cases. In short, most studies susceptible of rigorous hypothesis-testing have dealt with surface traits which psychodynamically oriented people sometimes regard with misgiving as an oversimplified correlating of symptoms, rather than a study of the persistent inner roots of behavior. Conversely, not many "depth" studies of personality structure have provided systematic measurements of child and family characteristics in samples large enough to allow mathematical tests of significance; and where such work has been done, it has generally been on "problem" samples, such as delinquents (3).

METHOD

Sample

In the course of an eight-year longitudinal study of adolescent character development, data became available that offered an opportunity to measure certain key dimensions of personality structure, and to relate them to fairly universal dimensions of family experience in a quantitative way. The sample population represented a relatively even cross section of all the children born in the year 1933 in a sociologically typical midwestern town, known here as "Prairie City." The intensity of the study limited the sample to 34 children (half girls and half boys) and their 34 families, but at least some tests could be made of some of the salient propositions now current about the influence of family constellations on dimensions of personality structure. Indirectly, the data permit a test of the validity of such constructs as "ego strength," "superego strength," and the like.

The population consisted of 34 children, selected from the 120 children living in Prairie City in 1943 who had been born there ten years before. The original selection was based on a composite assessment of personal adjustment, comprised of teacher-ratings, test scores and clinician ratings. A pilot study indicated that the 34 children were quite evenly distributed along a scale from the best adjusted to the most poorly adjusted in the total group of 120. (As a bit of parallel evidence, they were evenly distributed, also, from top to bottom of the total group on a composite rating of moral character reputation, based on peer and teacher ratings.)

The Research Data

Each child was intensively observed, tested, and otherwise studied from age 10 to age 18. There were annual interviews by resident field workers with the child, his or her parents, and with friends and other informants. Sociometric tests of varied kinds were obtained annually, as were projectively analyzable essays. The Rorschach was administered at 13 and 14; the TAT at 13, 15, and 17; and a Sentence Completion test at 16. Numerous intelligence tests were repeated each year or two. A variety of self-rating questionnaires on values, attitudes, personal adjustment, family adjustment, etc. were obtained annually. Teacher ratings of several kinds, anthropometric measures and certain other data were also gathered.

Procedure

In 1946 a research staff of 12, under the direction of William E. Henry and the late Caroline Tryon, was engaged in a study of the adolescent role behavior of the children in the sample. Interested in the developmental backgrounds of the subjects, the staff, among other things, rated the children's families on eight aspects of their emotional relationships and child-rearing practices. These data were subsequently filed away and remained unexamined until the work of the 1950 conference had been completed.

In 1949–50, a new staff of ten, under the direction of Robert J. Havighurst and the writer, conducted a study of the character development of the same 34 adolescents. An intensive case study was made of each child. Among other things, the staff made ratings (as of the

subjects' sixteenth year) on some 30-odd personality variables and, also, on two aspects of parental discipline (4). The estimated—Spearman-Brown—reliability of the judges' pooled-average ratings for any subject on any trait was .96.

Thus, two independent sets of data were available: one on the characteristics of the children's families, and the other on the personalities of the children. (The two additional family ratings by the 1950 staff were retained for analysis, although these were not strictly independent of the personality ratings, being made by the same judges.) A quantitative comparison of the measures of family dynamics with the personality dynamics of the children supplies the basis for the present report.

RESULTS AND DISCUSSION

Dimensions of Family Interaction

The ten family variables were intercorrelated and factor analyzed. Four trait-clusters were defined, with the following composition:[1]

F1 Consistency in Family Life
- Regularity of home life
- Consistency of parental control
- Common family participation in activities

F2 Democracy–Autocracy
- Child's degree of sharing in family decisions

F3 Mutual Trust and Approval
- Good father–mother relations
- Parental trust and faith in the child
- Child's willingness to share confidences with parents
- Parental approval of the child
- Parental approval of child's peer activities

F4 Severity–Leniency
- Severity of parental control

Dimensions of Adolescent Personality Structure

Twenty-nine of the personality variables rated by the 1950 conference were intercorrelated and factor analyzed (4). Six distinct trait clusters emerged as follows (the traits in each cluster are listed in descending order of their factor-loadings):

P1 Ego Strength[2]
- Emotional maturity
- Locus of concern (egocentric to altruistic)
- Internal consistency (roughly, personality integration)
- Rationality of behavior
- Accuracy of self-perception
- Accuracy in assigning responsibility (to self or others)
- Accuracy of social observation
- Functioning intelligence (largely, test-IQ)
- Insight into others' motives
- Autonomy (inner-directed)
- (Positive) Relations with the same-sex peers
- Positive feeling toward mother

P2 Superego Strength
- Superego strength (presence of an effectively behavior-guiding conscience)

P3 Willing Social Conformity
- Overt conformity to the (conference defined) moral code
- Emotional stability (largely, overt conformity to expected, "controlled" behavior)
- Range of moral horizon (the range of people and groups toward whom *S* is overtly moral)
- Overt acceptance of mother's expectations
- Absence of overt hostility
- (Positive) Relations with opposite-sex peers

P4 Spontaneity
- Identity of behavior with impulse
- Positive feeling toward father
- Empathy (with others' feelings)
- Positive feeling toward same-sex peers
- (Lack of) Guilt about overt behavior

P5 Friendliness
- Absence of covert hostility

P6 Hostility–Guilt Complex
- Guilt about inner impulses
- Inner feelings of hostility toward mother
- (Lack of) Overt acceptance of father's expectations
- (Lack of) Inner liking for opposite-sex peers

Relationships of Personality Dimensions to Family Characteristics

In order to compare the two sets of data, each child's standard score on each of the six personality dimensions was computed. Similarly, a standard score was derived for each family on each of the four family dimensions. It was thus possible to correlate the children's scores on the personality dimensions with the characteristics of their families.

From Table 1, the significant correlations suggest the following relationships, presumably causal in nature, between family experience and personality:

Ego Strength seems largely produced by Trust (.73) and Consistency (.56).

Social Conformity also seems importantly influenced by Trust (.60) and Consistency (.53).

Superego Strength seems influenced by Consistency (.50) and perhaps Trust (.33).

[1] See Table 1 for intercorrelations of these four trait clusters. These were factorially distinct clusters, not "pure," uncorrelated factors. See appendix of *Psychology of Moral Character* (4) for the rationale for this use of what Stephenson calls "mixed" factor types or clusters.

[2] See Table 1 for intercorrelations of these six trait clusters. The factoring was done in order to reduce the original, overlarge set of variables to a small number of more stable variables, with the result that six distinctly separate traits or trait clusters were found, defined by a three-factor system.

TABLE 1
CORRELATIONS OF FAMILY AND PERSONALITY VECTORS

Variable	F1	F2	F3	F4	P1	P2	P3	P4	P5	P6
F1 Consistency in Family Life		−.01	.66**	.37*	.56**	.50**	.53**	.05	.19	−.10
F2 Democracy–Autocracy			.53**	−.65**	.43*	−.07	.16	.36*	.33	−.40*
F3 Mutual Trust & Approval				−.22	.74**	.33	.60**	.27	.44**	−.40*
F4 Parental Severity					−.16	.26	−.08	−.38*	−.38*	.40*
P1 Ego Strength						.48**	.73**	.56**	.43*	−.33
P2 Superego Strength							.66**	−.24	.02	−.23
P3 Social Conformity								.11	.48**	−.37*
P4 Spontaneity									.39*	−.28
P5 Friendliness										−.75**
P6 Hostility–Guilt Complex										

* Significant at .05 level.
** Significant at .01 level.

Spontaneity seems influenced by Democracy (.36) and perhaps Trust (.27), plus a lack of Severity (−.38).

Friendliness seems most influenced by Trust (.44) and Democracy (.33).

The Hostility-Guilt Complex seems influenced by Severity (.40) combined with a lack of Trust (−.40) and of Democracy (−.40).

The over-all pattern shown in Table 1, suggests that mutual trust among family members (F3) is rather closely related (.66) to consistency in family life (F1). It might be inferred from the correlation pattern that these two family characteristics, working together, are responsible for the development of healthy *ego strength* (P1) and willing *social conformity* (P3). That is, a child's emotional maturity, personal integration, autonomy, rationality of behavior, and his willingness to adapt to society's expectations, appear to be directly related to the degree of consistency, mutual trust, and mutual approval he experiences within his family.

Superego strength (P2) is not unrelated (.48) to ego strength (P1), but insofar as it is a separate entity, it appears to be the product of a somewhat different pattern of these family forces. It seems mainly to come from a regular, consistent family life. Superego strength is not at all the same thing as the hostility—guilt complex (P6); indeed, they correlate slightly negatively, −.23. Moreover, superego strength is unrelated to the generalized attitude of friendliness (−.02). One might therefore think that superego strength would be associated with severe parental autocracy (F2, F4); but it is not. Judging from the case studies, the reason appears to be that while some children with strong, *rigid* superegos do tend to come from severe, autocratic families, there are other children who have equally strong internal principles but who test and apply them rationally. These latter children came from families that have been quite democratic and unsevere. Other data from the study suggest the existence of these two different *kinds* of strong superego, each of which has a different relationship with the democratic-autocratic family dimension. Spontaneity (P4) and ego strength (P1) are moderately related. Both characteristics appear to be favored by a lenient, democratic family atmosphere which actively promotes a self-accepting, good-natured spontaneity in the child.

Friendliness is significantly related to the degree of mutual trust and affection within the family, whether or not there is stable consistency in it. There are some friendly children who have weak egos and weak superegos. These come (in the present study, at least) from inconsistent, irregular families who follow a policy that might better be called laissez-faire than "democratic," though the parents show an uncritically acceptant attitude toward the children in a desultory, rather inattentive way.

The hostility-guilt complex (P6) might readily be called a hostile but dependent, unresolved Oedipal complex, upon examina-

tion of the variables that comprise it (see above). Not surprisingly, it tends to occur in children whose families are severe, autocratic but unloving and inconsistent.

SUMMARY AND CONCLUSIONS

The purpose of the study was to test and measure the relationships between key elements in personality structure and certain characteristics of family emotional and regulatory patterns. Thirty-four adolescents were selected as a representative cross section of all children born in Prairie City in 1933. They were interviewed and diversely tested each year from age 10 to age 18.

Independent ratings were made on the families, and on the children, by two separate research staffs, following intensive individual case studies. The family ratings were intercorrelated and factor analyzed, yielding four dimensions of family interaction: Consistency, Democracy, Mutual Trust and Approval, and Parental Severity. The personality ratings, similarly analyzed, yielded six dimensions of personality: Ego Strength, Superego Strength, Willing Social Conformity, Spontaneity, Friendliness, and a Hostility—Guilt Complex.

Scores for each child were then computed on each personality dimension and these were correlated with the corresponding families' scores on each family dimension. The intercorrelations were analyzed, to discover general patterns of relationship between family experience and personality.

In this sample of adolescents, covering the full range of personal adjustment and moral responsibility, there proved to be a significant pattern of relationships between family experience and personality. Specifically:

1. Ego strength occurred in association with a family life which was characterized by stable consistency and warm, mutual trust and approval between the parents and between parents and child.

2. Superego strength was partially related to ego strength, but was chiefly related to the regularity and consistency of family life. It was *not* systematically related to severely autocratic rearing; or, at least, there appeared to be two different kinds of superego, produced in two different ways: (*a*) a strong, rigid, compartmentalized superego created by sternly autocratic rearing; and (*b*) a strong superego which was closely knit with ego functions, open to rational appraisal, and created by consistent, democratic, nonsevere rearing in a trustful, approving family. This interpretation rests, of course, on the qualitative case studies.

3. Generalized friendliness and spontaneity appear to be allied and to be associated with a lenient, democratic family atmosphere. Probably because the family background may or may not be consistent, these personality characteristics therefore may or may not be allied with ego strength and superego strength.

4. The hostility—guilt complex might reasonably be considered a hostile but dependent, unresolved Oedipal complex. It tends to occur in association with a severely autocratic, untrusting, and disapproving family.

In general, insofar as the findings of this study bear on the concepts of present day theory about personality development, they tend to corroborate that theory. These adolescents' personality characteristics proved to be significantly related to the emotional relationships and the disciplinary patterns which they experienced in living with their parents.

REFERENCES

1. SEARS, R. R., WHITING, J. W. M., NOWLIS, E., & SEARS, PAULINE S. Some child rearing antecedents of aggressions and dependence in young children. *Genet. Psychol. Monogr.*, 1953, **47**, 135–234.
2. WHITING, J. W. M., & CHILD, I. L. *Child training and personality*. Yale Univer. Press, 1953.
3. GLUECK, S., & GLUECK, ELEANOR. *Unraveling juvenile delinquency*. Harvard Univer. Press, 1950.
4. PECK, R. F. *The psychology of moral character*. Unpublished doctoral dissertation, Univer. of Chicago, 1951. (Microfilm)

Received August 26, 1956.

CASE REPORT

SOME DYNAMIC ASPECTS OF POSTHYPNOTIC COMPLIANCE[1]

MILTON J. ROSENBERG AND CHARLES W. GARDNER[2]
Yale University

In this paper we suggest, and offer case-record evidence in support of, a point of view that attempts to account for interindividual and intraindividual variability in compliance with posthypnotic suggestions. This point of view assumes that "depth" of hypnosis (if at all a meaningful concept) is neither the only nor the primary relatum of such compliance. Rather, the set of factors considered in this paper has to do with the content and phenomenal symbolic meaning to the subject of the posthypnotic suggestions that are directed to him.

THEORY AND HYPOTHESES[3]

It will serve our purpose to first regard the typical motives of the hypnotic subject and then to examine certain aspects of his relationship with the hypnotist. Our theoretical approach is based in part on a number of psychoanalytic contributions to the study of hypnotic phenomena (1, 2, 3, 4, 6, 7, 8).

The Subject

The public motivation of the experimental hypnotic subject (*S*) is for money, if any is offered, for "satisfaction of curiosity," and for association with a person of relatively greater status and power. These are derivatives of, and are overshadowed by, a need for a passive-receptive relationship with a person endowed with pseudomagical attributes (6). The hypnotic *S* seeks a new person in response to the felt threat of loss of, or disappointment by, some significant figure in his life. That he is so motivated is often made clear through his expression of feelings of isolation, unrest, dissatisfaction, or even depression (7, 8). In the quest for a secure relationship, the *S* is often ready to forego some of his gratifications and his own untrusted strength, in exchange for the illusion of strength in the new object. He seeks to use the hypnotist as a partner in the acting out of a fantasy of oral fulfillment, ultimately a fantasy of the reunion of self with a parental image. At the same time he needs to use the hypnotist as a safe link to reality in this regressive dramatization. His exhibitionistic needs, too, are satisfied in the hypnotic relationship; particularly through the opportunity to relate exotic or remote material and to perform unusual hypnotic and posthypnotic tasks. Still other needs may be brought to the situation by the *S*, and these can generally be detected in his particular preconceptions about the nature of hypnosis.

The Trance

The "trance" is achieved ostensibly through the hypnotist's use of a formula emphasizing rhythmic repetition, monotony, and suggestions to relax (as well as other more specific suggestions) and by the *S*'s attempt to relinquish ostensible control or power and to submit himself to the hypnotist.

By means of the great diminution of sensory input and of channels of reality verification (through immobility, eye closure, etc.), to which the only exception is contact with the voice of the hypnotist, there occurs a fading or weakening of the boundaries of the self (4). The *S*'s need for magical participation with the hypnotist produces collaboration with this change, culminating in a partial merging of the images of the self and of a "sleep-giving" hypnotist.

In general, the contents of hypnotic imagery[4] and associations to it point to an underlying fantasy of oral or auditory incorporation that facilitates regression to a state of "fusion" between the images of the self and a powerful other (in this case the hypnotist).

Danger Points

Each hypnotic experience is unique, but most *S*s encounter certain anxiety-arousing "danger points." Initially, danger is felt in the terms of pretrance misgivings about surrendering control, and about the propriety of what the hypnotist may require.

[1] The authors wish to thank two of their colleagues, Roy Schafer and Irving Schneider, for their helpful evaluations of the original manuscript.

[2] Charles W. Gardner is in the Department of Psychiatry, Yale University.

[3] The theoretical approach proposed here is intended to apply primarily to the situation in which hypnotic experimentation is being done with volunteer *S*s. One would not expect an exact correspondence with the situation in hypnotherapy which is undertaken with *ostensibly* different motivations.

[4] We are indebted to G. O. Morris for unpublished observations on the function of imagery in hypnosis.

Somewhat later in the process, anxiety about compliance appears in images or thoughts of sexual attack, mutilation, or castration. The more profound dangers accompanying further progress in hypnosis are the fears associated with orality; fears of extinction, of being swallowed up or bitten, of swallowing up the hypnotist or destroying him, and fears of aloneness and loss of the self or of others.

The *S*'s past history and current psychological state determine which, if any, dangers he experiences and how he will respond to these dangers. The anxiety aroused in these situations can lead to a reversal or "lightening" of the hypnotic process; or the *S* may maintain himself in hypnosis through the use of characteristic mechanisms aimed at inner avoidance of anxiety and at manipulation of the hypnotist's behavior or expectations. Improvement, with practice, in hypnosis and in following hypnotic suggestions, represents the gradual mastery of the anxiety arising in the danger situations and is achieved through rehearsal and stepwise resolution of the conflict.

Posthypnotic Suggestion

Posthypnotic suggestion introduces new and additional difficulties. It requires that the *S* continue the identification with the hypnotist while in a "waking condition." In this situation, two contradictory states must coexist; a state of identification with a gratifying, somniferent hypnotist; and an awareness of a real manipulating hypnotist. Another difficulty is that the posthypnotic suggestion itself can complicate the hypnotic relationship in various ways. It may, for example, put defenses against repressed impulses in jeopardy, as when the behavior called for would reverse a strongly reinforced reaction formation.

Typically, the presentation of a posthypnotic suggestion is accompanied by the suggestion that the *S* will have a temporary amnesia for all, or for a portion, of the hypnotic session. This suggestion in itself will frequently reawaken all the reluctance to surrender control that was first experienced in response to the original achievement of hypnotic trance.

Despite these various difficulties, many *S*s (even some who have not been fully successful in ordinary "trance") do manage to comply with posthypnotic suggestions and to manifest amnesia for them. But, as we have noted, the success of such compliance does frequently vary from *S* to *S* and, with regard to a single *S*, from one posthypnotic suggestion to another. It is in the light of this observation that we have formulated the question to which the remainder of this article is addressed: *what processes guide and govern the hypnotic subject's acceptance of, and adaptation to, specific posthypnotic suggestions?*

In the present experiment, the attempt was made to produce, through posthypnotic suggestion, changes in the *S*'s *attitudes* towards certain social issues or social objects. Such responses are not as easy to change as are more ego-peripheral behaviors. However, it is this very fact which makes them particularly useful in research that seeks to uncover the dynamics of posthypnotic compliance. For where change is difficult to achieve, the experimental hypnotic *S* will, through his motivated *striving* to achieve the change, reveal more about the dynamics of such adaptation than he would in situations where the suggested change is less challenging. Thus we would contend that the dynamic processes posited in the hypotheses stated immediately below operate not only in the case of posthypnotic suggestions of attitude change, but rather that they pertain to any posthypnotic suggestion that arouses some degree, however small, of resistance in the *S*.

Hypotheses on Posthypnotic Compliance

Hypothesis I. Compliance with a posthypnotic suggestion varies as a positive function of the extent to which the subject can interpret and manipulate its content into an acceptable form. This is accomplished through the employment of dynamic processes set into motion during hypnosis and used by the subject to maintain himself in hypnosis.

Underlying this hypothesis is the view that the posthypnotic suggestion, being an attempted manipulation or intrusion by the hypnotist, will be responded to in the context of the already existing atmosphere of the specific hypnotic relationship. A further implication of this hypothesis is that the fate of the posthypnotic suggestion will depend on its usefulness to the *S* in the dynamic state in which he finds himself as he strives to ward off the anxiety and other disruptions associated with his hypnotic experience.

The "dynamic processes set into motion during hypnosis" can be considered, for present purposes, to fall into two classes. These are: first, *the mechanisms* used by the *S* to deal with the imminent emergence of anxiety in the induction and trance; and second, the specific mixture of *affective reactions* generated toward the hypnotist and his aims throughout the relationship and crystallized during the trance.

In the first group, some of the commonly found mechanisms are: *denial*, initially of the hypnotist's presence or influence, but capable of extension beyond these to other perceptions; *projection*, involving for instance, the attribution to the hypnotist of unacceptable sadistic impulses; forms of *rationalization*, especially explanations of com-

pliance which disregard the passive aims of the *S*; *intellectual distancing*, such as the device of adopting a detached observer stance toward one's own inner processes; *play*, involving a heightened emphasis on the make-believe and experimental nature of the hypnotic performance; *turning passivity into activity*, and thus mastering a traumatic situation by actively reproducing it upon the self. The ultimate device called into action if hypnosis is successful, is *identification* of the self with the image of a "good" (protective) hypnotist, or a "bad" (aggressive) hypnotist. The device of identification, while facilitating hypnosis, also satisfies one of the basic needs of the successful hypnotic *S*.

The mechanisms that have been listed do not necessarily operate singly and separately. They may combine and fragment; for instance, the illusory reversal of roles often reported by hypnotic *S*s depends, for its effectiveness, upon denial, projection, and identification, all of which may serve in mastering the threat of passive submission.

Hypothesis I suggests that the posthypnotic suggestion will succeed in the measure that its content allows the *S* to organize, interpret, and execute it in a manner consistent with and guided by the mechanism, or mechanisms, that he regularly uses to overcome the danger points of hypnosis. It also points to the significance of a second group of dynamic changes: the *S*'s more enduring reactions to the hypnotist and his assumed aims, and the affects mobilized by these reactions. For example, the hypnotist may be reacted to as a frustrating, gratifying, unfriendly, protecting, or aloof person. Thus an *S* smarting under what he feels to be arbitrary domination or ruthlessness by the hypnotist may respond to the posthypnotic suggestion in ways determined by his affect. He may view the posthypnotic suggestion as a further assault, another justification for resentment, or another opportunity for masochistic submission. According to Hypothesis I, the chances for successful performance of the posthypnotic suggestion will be enhanced if its content lends itself to organization and interpretation that are consistent with the affects mobilized in the *S* by his experience of the hypnotic relationship.

Any one *S* will employ one or more mechanisms to maintain himself in hypnosis and will experience some out of the spectrum of possible reactions to the hypnotist; it is the potential consonance of these mechanisms and reactions on the one hand with the subject matter of the posthypnotic suggestion on the other hand that will, according to our Hypothesis I, determine whether and to what degree the *S* complies with the suggestion.

Hypothesis II. Where a posthypnotic suggestion activates a continuing inner conflict, compliance with that suggestion varies as a function of the degree to which it facilitates safe discharge of the warded-off drive (or drives).

Virtually any posthypnotic suggestion that directs an experimental *S* to perform some response which under nonhypnotic conditions would be ego-alien, may thereby tend to activate some related area (or areas) of conflict. In our own experiment, the social attitude picked for change frequently represents the result of defensive restrictions against certain unacceptable impulses. Suggesting the reversal of such an attitude increases the pressure of the warded-off material (through the implied permission granted by the suggestion) and may thereby require added defensive operations. When the motivation for the maintenance of hypnosis is not strong these defensive operations may interfere with the trance or work against the execution of the posthypnotic suggestion. But more typically these defensive operations are affected and limited by the regressive concomitant of hypnosis.

Thus what is predicted in Hypothesis II is that in the situation in which the posthypnotic suggestion mobilizes an area of conflict, the typical hypnotic *S* will the more readily accept and execute the suggestion to the extent that it permits him safely to indulge and express previously warded-off but persistent impulses. This "safe" indulgence is made possible not only by the regressive state characteristic of hypnosis but is also facilitated by the *S*'s identification with the hypnotist whereby the hypnotist becomes a new institution within the *S*'s personality; as such, the hypnotist is experienced as a force which, *through* the posthypnotic suggestion, has the power to legitimize impulses which were previously experienced by the *S* as illegitimate or as endangering.

It is the impression of the authors that in situations in which the emotional problems and defensive tendencies of the hypnotic *S* can be studied, one or more of the processes postulated in our hypotheses will be found to govern the *S*'s success and style in complying with posthypnotic suggestions. In the study reported in the remainder of this paper, the available data permitted at least a partial test of the adequacy of our theoretical account.

Method

Eleven experimental *S*s were selected on the basis of their having demonstrated good potential hypnotizability. These *S*s were tested (with a questionnaire device) on their attitudes toward various social issues. Each *S* was then rehearsed in hypnosis to the point where he could carry out posthypnotic suggestions with amnesia for these suggestions. The number of sessions required to

reach this criterion level varied from three to eight separate hours of hypnosis.

The *S*s were then individually run through an experimental session which consisted of a number of steps. First, each *S* took a test[5] designed to measure the *cognitive structure* of one of the attitudes that he had already expressed on the attitude questionnaire.

On this "cognitive structure test," developed earlier by one of the authors (5), the *S* works with 31 so-called "value cards." Upon each card is printed a different value term (i.e., a state, condition, behavior, or attribute that may be positively or negatively, strongly or weakly, valued by the *S*). Examples are: "all human beings having equal rights"; "people being well educated"; "making one's own decisions"; "giving expression to feelings of anger or hostility"; "complying with the wishes of persons in authority." The *S* first judges each of the 31 values in terms of its importance to him, using a scale with a range of 21 points. This scale runs from "gives me maximum dissatisfaction" (Category −10) through "gives me neither satisfaction nor dissatisfaction" (Category 0) to "gives me maximum satisfaction" (Category +10). The intervening categories represent the intervening degrees of dissatisfaction or satisfaction. In a second part of the test, he is required to judge these same value items in terms of whether, and to what extent, he thinks each value will be *attained* or *blocked* as a consequence of the attitude object. On this task he uses an 11-point scale running from "extreme blocking of the value through the influence of the attitude object" (Category −5) through "the value is neither attained nor blocked through the influence of the attitude object" (Category 0) to "extreme attainment of the value through the influence of the attitude object" (Category +5).

Thus after testing there are available for each value term, the *S*'s judgment of its importance to him as a value and his judgment of whether and to what extent that value will, as he sees it, be attained or blocked through the influence of the attitude object. These two judgments are algebraically multiplied for each value item respectively. By summing all the products, a single algebraic quantity is obtained representing the total import of the *S*'s pattern of beliefs about the influence of the attitude object upon attaining or blocking various states that he values to differing degrees. This "index of cognitive structure" was computed for each *S*.

Upon completion of the cognitive structure test, the *S* was hypnotized to a stable level. He was then given a suggestion of posthypnotic affect-reversal with regard to the attitude object upon which he had been tested for associated cognitive structure. For example, one of the *S*s who had expressed, on the attitude questionnaire, strong opposition to "living in Los Angeles" was given exactly the following hypnotic instructions:

> When you awake you will feel very much in favor of living in the city of Los Angeles. The mere idea of living in Los Angeles makes you feel happy and exhilarated. Although you will not remember this suggestion having been made, it will strongly influence your feelings after you have awakened. Only when the signal to remember is given will you be able to remember and only then will your feelings revert to normal.

The posthypnotic suggestion, then, was simply of a change in affect on some attitude issue. Nothing whatsoever was suggested about how the person would *think* or what he would *believe* on that issue. The specific experimental prediction was that *S*s who complied with the posthypnotic suggestion of affect-reversal would undergo a process of spontaneous cognitive reorganization modifying their beliefs about the value-attaining and value-blocking potentialities of the attitude object so as to rationalize the affect-reversal.

To test this prediction, the *S*, after being brought out of hypnotic trance, was retested with the attitude questionnaire device (to determine whether the posthypnotic suggestion of affect-reversal was being complied with) and then with the test of cognitive structure (to determine whether he had changed his beliefs in a way which served to rationalize the affective change).[6] Upon completion of these tests the *S* was interviewed on the same attitude issue and also on his memory for, and feelings about, the hypnotic session. Finally the amnesia-removing cue was presented and after the *S* had achieved recall for the suggestion, the nature of the experiment was explained to him.

The full results of this experiment will be reported elsewhere.[7] It will suffice here to report that the original prediction was fully borne out and that the experimental *S*s showed significantly more cognitive reorganization ($p < .005$) than a control group not exposed to hypnotic suggestion.

Although it was not possible to analyze all of the resulting data in terms of the personality dynamics of the *S*s, with four of the *S*s circumstances were such that a large number of additional

[5] All tests were administered by a person other than the *S*'s hypnotist.

[6] The full "cognitive structure test" records for all of the *S*s are reported in an addendum that has been deposited with the American Documentation Institute. Order Document No. 5721, from ADI Auxiliary Publications Projects, Photo Duplication Service, Library of Congress, Washington 25, D. C., remitting $1.25 for 35 mm. microfilm or $1.25 for 6 by 8 in. photocopies.

[7] In a forthcoming publication by one of the authors (M. J. Rosenberg).

nterviews could be conducted along psychiatric lines. With each *S* these interviews were focused on his conflicts, mechanisms, and personality-structure generally and also on his emotional reactions in, and defensive reactions to, the hypnotic process. Only when the dynamic interpretations drawn from these interviews were completed were they brought to bear upon analyzing the *S*'s performance in response to the specific posthypnotic suggestion he had received. In each case, the data on the *S*'s regnant conflicts and defense mechanism priorities and their influence upon his experience of hypnosis helped to illuminate the dynamics of his acceptance of, and adaptation to, the posthypnotic suggestion; and in each of these cases the resulting findings seem clearly to exemplify and confirm the hypotheses on posthypnotic suggestion that have been reported above. The next section of this paper is devoted to the presentation of these four cases.[8]

Subject D

Subject D was a young single woman, living alone in the city, and temporarily unemployed. She had been born and raised in Europe and at the time of the experiment had been in the United States for about five years. While she had an educational background and skills that suited her for an independent career, she had been employed in a field remote from her own. Relatively isolated, she was somewhat unapproachable, and heightened this defense by keeping as a pet an intimidating-looking dog. It was upon this dog that most of her affection was spent, and arrangements for her hypnotic sessions sometimes took second place to her providing care for her pet. She appeared to be in a chronically dejected mood, feeling sad, lonely, and suffering feelings of discouragement and emptiness. She had characteristically a wistful, pathetic manner, and elicited from those around her reactions of sympathy and gentleness, along with an awareness that she was still aloof and quite uncommunicative about herself.

The major historical items in her life are not essential here, but it is of interest to note that she came for hypnosis at a crucial juncture, and used it in a specific way. It was while working as a hypnotic *S* that she applied for, and got, a job somewhat closer to her interests; and it was during the course of the experiment that she arranged to do work, on a commission, for which she was really trained and which promised some success. Once this commissioned work was undertaken, she found herself unable to continue as a hypnotic *S* in a later experiment, and appeared to have made a more stable and more satisfying adjustment to her needs.

Among her original expectations about hypnosis was the notion that it would provide a relaxed, restful oblivion; something she yearned for but felt she had never been able to achieve. At the first hypnotic attempt, she went readily and "deeply" into hypnosis, exhibiting utter docility, and appearing to be a "perfect subject." This performance was not only gratifying to the hypnotist, but left the *S* with a feeling of profound exhilaration and satisfaction. Throughout the experiment there was a marked contrast between her rather apathetic, tired appearance before a hypnotic session and her comfortable, pleased look afterwards. She became enthusiastic about the experiment and practiced hypnosis upon herself at home. Throughout our work with her she would happily and spontaneously report upon her progress in inducing various hypnotic phenomena in herself. She understood her attempts at self-hypnosis as a kind of psychotherapy, a search for which was a subsidiary motive in her volunteering, in addition to her basic need for satisfying human contact.

Corresponding to her insulated mode of existence, the *S* displayed no strong sentiments on current or basic human issues. Except for her interest in her pet, which was protective and coddling, and a not always satisfying relationship to her work, she claimed no strong fixed opinions or attitudes.

Course of hypnosis. The major maneuver used by this *S* to surmount the dangers of hypnosis was the turning of a passively-experienced process into one actively reproduced by herself. That is, she reported consistently that it was not the hypnotist who gave her instructions under hypnosis, but herself. She exaggerated this not uncommon maneuver and carried it over into "waking" life, acting the part of the hypnotist to herself.

Although largely inarticulate about herself, she used the currency of her mood to communicate with the hypnotist. Thus her painful sadness at the outset, and her hypomanic state at the conclusion of a hypnotic session functioned as a plea and a reward to him. The appearance of submissiveness was enhanced by her frequent apologies for not "doing well enough," her fear that the experimenters would not be entirely pleased, and her eager promise to try to "do better" next time. In summary, she went to an extreme in almost abject compliance (against whose dangers she guarded herself by use of the mechanism described above) in exchange for the sensation of a consoling relationship with a powerful figure.

Adaptation to the posthypnotic suggestion. In this

[8] For each of the *S*s, the procedures that have been described were repeated for a second attitude area not dealt with in this paper. We were able to apply our interpretive procedures to the second attitude change for three of the four *S*s. In each of these cases the same individual dynamic patterns appeared as described here. With the fourth *S*, the change achieved was of too small a magnitude to enable analysis in depth.

case, the posthypnotic suggestion was concerned with the *S*'s attitude toward the "city-manager plan." Prior to the experimental session she had indicated, on the attitude questionnaire, a moderately opposed position. Her score on the test of cognitive structure was −118, representing the fact that she perceived many bad consequences to follow from the city-manager plan.

In the experimental hypnotic session she was given the following suggestion of affect-reversal:

> When you awake you will feel very much in favor of the city-manager plan. The mere idea of the city-manager plan makes you feel very approving and enthusiastic. Although you will not remember this suggestion having been made, it will strongly influence your feelings after you have awakened.

Immediately following the trance in which this suggestion was delivered, the *S* was retested on the attitude questionnarie and on the test of cognitive structure. On the attitude questionnaire she now indicated that she was extremely "in favor" of the city-manager plan. This was accompanied by a spectacular shift on the test for cognitive structure to a score of +311. The *S*'s amnesia for the suggestion of affect-change was equally complete and thorough-going.

Her comments on the change suggested that the same mechanisms were operating as she employed in her hypnotic trance, along the lines of our Hypothesis I. Thus, in the *S*'s view, the change was accomplished by a "redefinition" of just what a city-manager plan was. She reported that prior to the experimental session, she had assumed that "it was a thing that restricted you—more or less enforcing all sorts of things." But, as she explained it, after the session, "I think I changed my assumption a little bit—it seems as though it all kind of reversed—I think really I feel the same way about it—only I just have stressed the ... another angle of it." Now, after the experimental session, the city-manager plan was perceived as "something voluntary which the people themselves would want to do for the city—and have everybody represented. I think maybe the first time [before the experimental hypnotic session] I thought the people wouldn't be represented and it would just be kind of inflicted on them—city bosses would get together and make decisions and possibly decisions that affect their personal interest instead of the people's welfare."

Later in the same interview she said, "at first I thought it would be a type of control which I have seen which I heartily dislike—for instance in cities where they invite all sorts of industries to come regardless of what type it is. And then the second time I had in mind a sort of more democratic system in which people take into account each others interests." She reported having been unaware of the "democratic" definition of the city-manager plan prior to the posthypnotic suggestion. She attempted to explain this change in the meaning of "city-manager plan" on the basis of her having "whims," being changeable, and said of herself, "I'm not terribly consistent in the way I think."

The attitude-change suggestion, then, has been complied with by treating it precisely as she has treated the hypnotist's activity in hypnosis: by constructing a "redefinition" of the concept "city-manager plan" whereby it loses its authoritarian, mastering implications, and acquires a measure of voluntarism. This more generous view of the city-manager plan as "democratic" corresponds to her more comfortable acceptance of hypnosis when she can turn a passively experienced event into one actively reproduced by herself so that, in effect, her interests are observed and taken into account. This shift in the meaning of "city-manager plan" can also be seen as a carry-over of her enduring affective state in hypnosis; a state of a pleased and grateful feeling toward the hypnotist who has become, not a tyrant, but a kindly figure. In brief, the adaptation to the posthypnotic suggestion seems, in this *S*, to be a rather clear illustration of the operation of the dynamic process which is posited in Hypothesis I above.

Her changes on the test of cognitive structure add confirming detail to the interpretation that has just been given. For example, on the test before the experimental hypnotic session she judged the value item, "complying with the wishes of persons in authority," as being "dissatisfying" (rating of −8 on the value satisfaction dimension) and indicated that she believed the city-manager plan tended to foster this negative value. After the administration of the posthypnotic suggestion, this same value item, while still judged as one which is dissatisfying, is now seen by the *S* to be "blocked" and prevented by the city-manager plan. For this *S*, the value item of "complying with the wishes of persons in authority" is probably representative of the idea of loss of autonomy. By transforming the meaning of the city-manager plan into a concept stressing its "democratic" features, she is able to deny the element of compliance both with regard to the posthypnotic suggestion and with regard to hypnosis itself.

A similar change occurred with the value item, "having power and authority over people," which is considered by the *S* to be a dissatisfying thing and is seen as attained through the city-manager plan before the posthypnotic suggestion and as blocked by it after the posthypnotic suggestion.

The mechanism she has used—of becoming her own hypnotist and thus effecting a successful merging of a benevolent, sleep-giving image of the hypnotist with the self—is reflected in four

value items which, prior to the trance, were seen as blocked by the city-manager plan but after the trance are seen as fostered and attained by the city-manager plan. These four value items are "all human beings having equal rights"; "the open expression of disagreement between people"; "people of different backgrounds getting to know each other better"; "people having the right to participate in making decisions which will affect them." Her posthypnotic classification of these four values seems to bring to bear upon the task of rationalizing the posthypnotic suggestion, the rewarding closeness and equality that she has achieved in hypnosis through her mechanism of denying that hypnosis involves any domination.

In these and other aspects of her rationalization of the affective posthypnotic suggestion, the *S* gives further evidence that for her, acceptance of the posthypnotic suggestion is guided by the type of dynamic process that we have pointed to in Hypothesis I: the carry-over of the same mechanism to the posthypnotic suggestion as operates in defending the *S* against the danger points of hypnosis.

In view of the lack of substance of her fixed attitudes, Hypothesis II is not called into play. As she herself states, she generally has little investment in social questions and in fact may adopt or relinquish a point of view in order to please or conform with an expectation of someone else. Thus she appeared interested in the details of her attitude change, but did not feel strongly about the fact that she had changed. Rather she was more concerned about continuing work as a hypnotic *S* in a further experiment. Her greatest satisfaction after the experimental session appeared to be that she had performed well and had pleased the hypnotist.

Subject F

Subject F was a 21-year-old single student who was in his last year of preparation for divinity school. Of the four *S*s whose hypnotic performances are analyzed in this paper, he is the one on whom we have the least amount of clinical material. But interviewing and observation were sufficient to suggest certain salient facts about his character structure and his techniques of defense. The overall picture seemed to be of a sado-masochistic person in whom the instinctual forces of destructiveness were tightly bound and rechannelled through a life pattern based on reaction-formation. While acknowledging that he had no true "religious faith" and felt considerable skepticism about the dogmas of his and other denominations, he saw the ministry as a type of work through which social good might be done. For him its ethical imperatives were the source of such validity as Christianity possessed.

During his years as a student he had held a part-time job as an attendant in an inpatient psychiatric hospital and found this work "fascinating." It would appear that in general he was able, through the mechanisms of denial and reaction-formation, to control his sadistic impulses by transforming them into an ideologized charitableness and "good will." These he expressed in part through his involvement in the functioning of the psychiatric institution; but in this setting his superordinate relationships with the emotionally ill allowed him, at the same time, a continued gratification (through the secret pleasures of "Schadenfreude") of the sadistic component.

His style of conversation, if expressive cues were ignored, was friendly and cooperative, sometimes diffident. But by the use of his voice and his tendency to produce long mid-sentence pauses during which he would scowl, examine the room and then, in muffled tones, continue his comment, he seemed to reveal a persisting suspiciousness and an emotional substratum of irritability.

From his first contact with the experimenter he expressed great interest in hypnosis and an eagerness to "see for myself what the hypnotic experience is like." He was able from the first session to accept a degree of hypnotic involvement that included the achievement of motor effects, anesthesia, compliance with posthypnotic suggestions and posthypnotic amnesia.

Course of hypnosis. During each of the training sessions, the *S* entered rapidly into the trance state. A common accompaniment of entry into trance was loud deep breathing and a heavy thrashing about of limbs and trunk. The posthypnotic suggestions were almost always carried out with a rather exaggerated compliance to the implications as well as to the letter of the instructions. Thus when told that a certain book would appear to be printed in German he shouted upon first opening its pages, "This is fantastic—it's in German." He then proceeded to read aloud a paragraph from this book with the experimenter looking on. His "reading" consisted of a rapid and fairly accurate sight translation, into German, of the English text.

In the discussions that followed each training session he would first manifest a sense of dullness and confusion, sometimes seeming rather stuporous. On one such occasion he commented that he was not sure whether he was or was not still in hypnosis. As his confusion wore off he would begin to raise questions about the nature of hypnosis. In doing this he seemed to be indirectly expressing and stilling his anxiety about the intense relationship into which he had allowed himself to enter.

The contents of the posthypnotic conversations and of the later interviews suggest strongly that for this *S* hypnosis was a very fulfilling sado-

masochistic encounter; one in which by total "submission" he achieved a rewarding identification with the hypnotist and thus partook of the hypnotist's magical powers of dominance, control, and omnipotence. His pressing needs for such sado-masochistic fulfillment were, however, almost totally inadmissable in terms of his conscious self-image and his usually effective employment of the mechanisms of denial and reaction-formation. The resulting, and potentially disruptive, conflict he usually handled by the expedient of dullness and inarticulateness in the period immediately following emergence from the trance state. In this way he was able to isolate the ego state obtaining during hypnosis and the motivational forces with which it was identified as though they referred to a different person; i.e., an amalgam of self and hypnotist. He thus avoided the requirement of integrating these unassimilable impulses with his usual self-image.

Adaptation to the posthypnotic suggestion. With this *S*, the posthypnotic suggestion of affect-reversal was focused on his attitude toward "Negroes moving into white neighborhoods." Prior to the experimental session he had indicated on the attitude questionnaire instrument a strong positive reaction to "Negroes being allowed to move into white neighborhoods." However, his total score on the test of his cognitive structure about the housing desegregation issue was +3. This score was based upon his holding separate beliefs that Negroes moving into white neighborhoods would have both positive and negative consequences. The discrepancy between his strong positive score on the test of attitudinal affect and his mixed and ambivalently organized cognitive structure suggests that he felt required (whether by the situation or by his own ego ideals) to represent himself as fully committed to the prodesegregation position, whereas he was in fact capable of thinking about the issue in a way which reflected some antagonism to the possibility of Negroes moving into white neighborhoods.

In the experimental hypnotic session, he was given the following suggestion of affect-reversal:

> When you awake you will feel great loathing and disgust toward the idea of Negroes being allowed to move into white neighborhoods. The mere idea of Negroes being allowed to move into white neighborhoods offends you and makes you feel quite angry. Although you will not remember this suggestion having been made it will strongly influence your feelings after you have awakened.

While listening to the instructions about his posthypnotic feeling on the issue of Negroes moving into white neighborhoods, the *S* emitted a series of rasping, growl-like sounds. These ceased after the presentation of the suggestion was completed. Immediately after emergence from the trance and before the readministration of the tests could begin, the *S* began to manifest various expressive signs of anger. He kept his lips tightly pressed together, his fists clenched and spoke with a clipped, stacatto rhythm.

On the second administration of the attitude questionnaire, he marked as his own the extreme negative position with regard to Negroes moving into white neighborhoods. In filling out the Negro desegregation item he continued to show high tension, held the pencil very firmly, rewrote his answer several times and then placed a number of asterisks next to it. As the testing continued, he frequently broke into spontaneous comments of hostility toward Negroes, either shouting or snarling most of these comments. For example, at one point during the cognitive structure test he said with considerable fury, "I'm prejudiced. I'm against Negroes moving into white neighborhoods. They should stay in their own dirty, sloppy run-down neighborhoods and not spread destruction into white neighborhoods." In the interview following the completion of the test procedures, the *S* continued to express hostility and condemnation toward Negroes moving into white neighborhoods.

Despite his frequent outbursts against Negroes, his behavior during the cognitive structure test was confused and disordered. It was as if, although expressing very strong anti-Negro affect, he was incapable of rationalizing that affect through the provision of arguments that justified it. The card-sorting test of cognitive structure, which he had taken and fully understood before the experimental hypnotic session, now seemed to confuse and overwhelm him. He completed it only by throwing the cards into categories in an almost haphazard fashion, all the while voicing the complaint that the test was confusing, senseless, and boring. This general avoidance of the rationalization task posed by the test of cognitive structure renders his specific card-sorting performance irrelevant for the present analysis. But the *fact* of this avoidance is of great interest in interpreting the *S*'s general adaptation to the posthypnotic suggestion.

It has already been suggested that this *S*'s easy entry into, and tolerance for, hypnosis is based upon the opportunity for gratification of his sadistic wishes. Through his identification with the hypnotist he is experiencing in hypnosis that sense of manipulative and potentially injurious sadistic control which he ordinarily transforms through denial and reaction-formation. Now in the experimental session he is told by the hypnotist that he will feel negatively toward Negroes moving into white neighborhoods. This suggestion gives him permission to enter into a virtually uncontrolled outburst of primitive rage, and to direct that rage at one of those objects which, by virtue of social

tradition, has always been available for attack if the attacker can tolerate acting out his own hostility. Such a tolerance for his own sadistic impulses is not normally possible for this *S*, but because of the release that hypnosis has already provided and because of his temporary assimilation of an image of a strong, sadistic, and responsible hypnotist, he is now free to enjoy, in comparative safety, the release of his sadistic impulses.[9]

Thus the dynamic process that seems to govern this *S*'s acceptance and execution of the posthypnotic suggestion of antidesegregation is the one encompassed in our second hypothesis. The posthypnotic suggestion is effective through its affording a comparatively safe abandonment of a basic reaction-formation; an undoing and temporary cancellation of ego control which frees, for direct expression, a driving force of primary significance in the *S*'s emotional life.

The near violence with which this *S* expressed the anti-Negro-desegregation affect is probably of significance in accounting for his inability to rationalize this affective change through his responses on the cognitive structure test. We have seen that the *S* maintains a posthypnotic pattern of dullness and confusion that enables him to keep the satisfactions of hypnosis discrete from his usual picture of himself. Similarly, the strength of his reaction-formation to his sadistic impulses requires that he keep the outburst of anti-Negro feeling out of connection with his usual ego state. Using the same mechanism of functional stupidity, he finds the cognitive-structure test confusing and himself dull and muddled. Thus he avoids having to offer an argumentative defense of the anti-Negro affect and this avoidance enables him to keep his expression of that affect out of association with his usual patterns of thought. In this aspect of his adaptation to the posthypnotic suggestion we seem to have the sort of dynamic event dealt with in Hypothesis I, namely: an execution of the posthypnotic suggestion through the employment of the kind of mechanism which is regularly used by the *S* to overcome the danger-points of hypnosis and to maintain himself in the hypnotic relationship.

After the testing was completed, the experimenter had a further discussion with the *S* before giving the amnesia-removing cue. Despite numerous questions about his reasons for feeling so strongly against Negroes moving into white neighborhoods, the *S* could not produce any coherent rationalizations except in one highly interesting exchange which is given here:

> Hypnotist: What do you have against Negroes moving into white neighborhoods?
>
> Subject: What do I have against it?
>
> H.: Yes.
>
> *S.*: Oh for heavens sake. The evaluation of property goes down, education is . . . is . . . and look what happens to our churches, they fall apart.
>
> H.: Do they really? It's something I don't know much about.
>
> *S.*: Why today the downtown church is a ruined church because of the Negroes moving into city areas, the people who . . .
>
> H.: The downtown—you mean on the Green?
>
> *S.*: Those aren't typical churches there. They have sense enough not to let Negroes come to those churches.

What is of particular interest in this excerpt is that virtually the only rationalization that the *S* could muster is one that involves a reference to "the church"; the very institution which in his conscious thought is the source, and effective instrument, of those moral imperatives which bind and transform his sadistic tendencies. Possibly he is working off some guilt over his aggressive outburst; guilt which has not been fully and successfully blocked by his restriction of the anti-Negro affect to a somewhat altered state of consciousness. Thus, to combat this guilt he improvises a perception of Negroes as threatening the agency that symbolizes those very moral standards which he has violated by expressing his sadistic impulses.

When the amnesia-removing cue was given, the *S* had complete recall for the original suggestion, immediately spoke of feeling great embarassment, and then went on to comment at length on how interesting it was that hypnosis could make one behave in a way so unlike one's "real self." All indications were to the effect that his final resolution of the incident was to once again seal over the sadistic impulses that he had indulged, so that his defensive structure remained easily as intact as it had been before the experimental session.

Subject J

The *S*, a married man with three children, was a conscientious, hard-working person of high intelligence, with a very good academic record behind him. He was attending professional school at an age somewhat older than usual, having had a career in a related field which did not satisfy him enough. He felt that his earlier area of work was only second-rate and, although worthwhile in itself, lacked the prestige of the profession he was now preparing to enter. His attendance at school had been at the cost of considerable sacrifice to him-

[9] It seems possible too, that the intensity of the *S*'s rage is due partially to a *displacement*, onto the content of the hypnotist's suggestion, of strong anger felt toward the hypnotist. Such anger, based upon the manipulative, intrusive aspect of the hypnotist, cannot be directly expressed if hypnosis is to be continued. But it can be displaced by virtue of easy associative paths, and by virtue of its congruence with underlying sadistic impulses, to the hypnotist's suggestion.

self and to his family; his wife had held two jobs at the same time to help support the family.

After a prolonged period of scraping by financially, Subject J was now able to proceed with further professional training because of outside help given him on the strength of his previous record. He felt himself to be on the verge of professional rise and fulfillment. He had aspired for many years to this kind of distinction that would set himself apart from and above his peers as an educated and competent person. He described himself as wanting "to be a big boy." Among the inner obstacles to his achieving this status were his awareness of some of his own dependent, "little boy" characteristics, and his considerable shame about his parents, their way of life, and his "messed up" home life as a child.

He had felt his family, in which he was the only child, to have been a source of shame and a secret to be kept from his friends. There were thoughts and ruminations about his family that he feared to reveal, feeling that were they known they would prejudice his chances of professional eminence and possibly ruin him or expose him to ridicule. In this connection he once said that now that both his parents were dead, it did him no good to talk about them since it would only gain him sympathy when what he wanted was advancement. He had sought this advancement steadily and successfully, but at this moment in his life found himself in the paradoxical situation of having professional success on the way and yet being dependent on financial help from others after having fought most of the way alone and unaided. He was proud of getting so far, of having pushed himself to his present level of competence and recognition.

However, his present family responsibilities seemed somewhat of a burden, and he complained at times of his irritability and curtness toward his children. There had been more serious sources for this irritability in the past; at one time, one of his children had required major surgery while his wife was ill and incapacitated. At the same time, his father, whom he had privately renounced and had always ambivalently loved, was in a terminal illness, and the *S* was required to bear part of the financial burden for that as well.

The *S* volunteered for hypnosis in response to a current need: to find a stronger and more respectable figure than his father with whom to participate and with whom to share, especially through work with professional and academic overtones. But at the same time, this father-substitute had to be debased and scorned so as to enable the *S* to assert his own independence and power. Illustrative of this conflict was the fact that very early in the experiment he endeavored to discover what degree the hypnotist held and was gratified when he found that, in his eyes, it did not carry as much prestige as the degree he himself was soon to receive. This ambivalent regard for the hypnotist was matched by some vacillation over attaching himself to him. He "confessed" during a later conversation that in preparation for one of the hypnotic sessions he had indulged in one of the habits that he detested in others: he had "braced" himself with a number of cocktails.

His characteristic way of dealing with affects was to avoid expressing them except in detached, intellectualized descriptions. Thus, he could describe suffering "dreadful fear," and itemize this with an accurate account of associated somatic manifestations while yet remaining visibly composed and even intrigued by his description. In most affairs he adopted the basic stance of a detached, almost-clinical observer, setting himself above the emotions he might feel. He had an eager readiness to introspect, to think of himself in psychological terms, to outdo the hypnotist's expectations in his candidness and open-mindedness about his inner life. Later interviews with the *S* confirmed the hypnotist's impression that by the clinical stance that he took toward himself, the *S* was attempting to reward the hypnotist and thus assure a prolongation of the relationship. At the same time, this aspect of the *S*'s behavior is a good example, of which there were many others, of his assuming some traits of the hypnotist by way of identification.

Course of hypnosis. The *S*'s fears about hypnosis concerned its possibly "spectacular" influence, its being somehow faintly improper, and its being a form of probing and "brain-picking" through which something about him might be revealed and even publicized. This last mentioned fear stood in distinction to his usual eagerness about his own "inner life." In addition to these there were fears corollary to his basic motivation for hypnosis: his need for association with a "big boy" of some prestige whom he can yet master; with whom he can be close, but still safe. These were fears of being dominated, overpowered and overcome.

The "danger points" in hypnosis were easily demarcated by the *S*. The first step in induction seemed frightening, even alarming at times, by virtue of its being a signal of his submissiveness. This experience of fear was often colored with ideas of too great intimacy, too sensual or sexual a closeness, and what he called a "threat to my masculine feeling." These fears (at one time he described hypnosis as being "like witchcraft, as though it smacks of evil") were dealt with by heterosexual fantasies of "sweater girls," and by his usual mechanism of robbing the situation of affect and converting it into something he himself was observing in an aloof way. The tests of hypnotic performance, eye closure for instance, were equally alarming to him, as were his experiences of "fall-

ing" and other changes in body feeling. At these times he would be aware of a lightening of his trance. Another way of dealing with these and similar danger situations was through identification with the hypnotist whom he sometimes saw as a protecting, feeding person. In his identification he *became* the hypnotist, lost any feeling of the hypnotist's nearness, retained only his words but not his voice as such. According to his report, it was "as though I thought the suggestions myself." Nevertheless, he often got the words wrong, and would occasionally carry out a suggestion much distorted, even caricatured or reversed. In this process he seemed to be trying to deny and limit the influence of the hypnotist. Despite this, he was an unusually adept *S* in most respects; the readiness with which he could be hypnotized eventually became his chief source of great anxiety.

His general attitude toward the experience of hypnotic trance is partially summarized in imagery he produced during one session in which he visualized a little boy being given bread and patted on the head and then assaulted cruelly by a "big monster." The fantasy ended with the death of the little boy under the wheels of a truck.

To recapitulate our analysis of this *S*'s adaptation to hypnosis: he seeks a father-replacement or substitute at a time when his usual reliance on his own great efforts for advancement is suspended by his receiving outside help. Distrustful of the intrinsic strength of his father-substitutes, he attempts to hold the hypnotist to him while at the same time belittling him; he scorns the hypnotist while at the same time making an identification with him. The dangers of the passive-dependent role in hypnosis are dealt with by such methods as avoidance of affect, the cultivation of a detached "observer" role, isolation, and a reassertion of his own personal strength.

Adaptation to the posthypnotic suggestion. The posthypnotic suggestion used with this *S* dealt with his attitude toward "the federal government providing comprehensive medical insurance for all persons." Before the experimental session, the *S* had characterized himself, on the attitude questionnaire, as being extremely in favor of federal medical insurance. His score on the test of cognitive structure was +195, representing the fact that almost all of his stated beliefs about federal medical insurance were of the positive variety.

In the experimental hypnotic session, the *S* was given the following suggestion of affect-reversal:

> When you awake you will feel great loathing and disgust toward the idea of the federal government providing comprehensive medical insurance for all persons. The very thought of this plan offends you and makes you angry. Although you will not remember this suggestion having been made, it will strongly influence your feelings after you have awakened.

Immediately following the hypnotic session the *S* was retested on the attitude questionnaire and the test of cognitive structure. As measured by the attitude questionnaire, the *S*'s attitude toward federal medical insurance was now given as *mildly* positive (this change representing a significant shift in the negative direction). A more extreme change was noted in the test of cognitive structure. From +195 before the posthypnotic suggestion, he moved to −58 after the posthypnotic suggestion, a score based on a cognitive structure containing beliefs of both positive and negative consequences following from federal medical insurance. In general, the posthypnotic suggestion in this case had the effect of rendering *ambivalent* an attitude that had previously been purely positive.

Immediately following the second testing, the hypnotist engaged the *S* in a conversation dealing with his feelings about federal medical insurance and about the experimental hypnotic session. Two particular themes recurred frequently during this conversation. The first of these themes was revealed in the *S*'s numerous references to the fact that it was only "the second time around" (i.e., after the experimental hypnotic session) that he came to realize that federal medical insurance was not a pure blessing but, rather, a mixed thing. The negative aspect of federal medical insurance which he now saw clearly and felt concerned about was that "it could infringe upon the rights and liberties of individuals." Frequent references to this idea marked it as the major area of change in the *S*'s posthypnotic thinking on the issue of federal medical insurance. It is of course noteworthy that in his earlier adaptation to hypnosis, the *S*'s major problem was to overcome his sense of danger at giving up the independence, autonomy and emotional distance that were so crucial to his general adjustment.

An interpretation that seems to us to be required in this case is that the *S* manages partially to interiorize the suggestion of negative affect toward federal medical insurance by finding that this object violates those very standards (of independence and autonomy) which he feels to be threatened by hypnosis itself. Thus in partially accepting the posthypnotic suggestion he is at the same time defending himself from the danger of loss-of-autonomy in hypnosis by rationalizing the suggested affect in terms of those standards which figure in his characteristic stance towards hypnosis. The process involved, then, would seem to be the one identified in Hypothesis I above.

A second theme that arose frequently in the discussion after the experimental hypnotic session was the *S*'s awareness of "feeling angry." As he

reported it, he had awakened from hypnosis "feeling a sort of clenched fist kind of anger." He remarked that perhaps this anger had been "set up" for him by the experimenter. Later when the amnesia was lifted, he felt the same sort of anger surge back when the experiment was explained to him. But though he described himself as feeling angry, he did not allow himself to express this feeling except by intellectually noting its presence as a subjective state.

A further complication in expressing his sense of anger at having been manipulated is revealed in the fact that, immediately after having mentioned feeling angry, he went on to express his "utter amazement" at the fact that his opinions had been changed by the hypnotist. This comment was followed by a repetition three times of the phrase, "Isn't that wonderful?"

What the *S* seems to be doing in this interlude after the lifting of the amnesia is to give continued expression of his mixed reaction to the hypnotist by way of commenting on the hypnotist's experiment. It seems to us possible that the *S*'s ambivalence toward the hypnotist and toward the hypnotic experience is reflected in his responding to the posthypnotic suggestion by developing an ambivalent attitude towards federal medical insurance. In other words, the content of the hypnotist's suggestion has been treated in the same way as the figure of the hypnotist; the content of the manipulation has been dealt with in the same way as the fact of manipulation. Such a process could be understood in terms of our Hypothesis I.

Both aspects of our interpretation seem to be borne out by a number of changes on the test of cognitive structure. For example, on the first test (before the experimental hypnotic session) he judged the value item, "people having the right to participate in making decisions which will affect them," as being neither satisfying nor dissatisfying to him (i.e., a rating of 0). He indicated further that, as he saw it, the institution of federal medical insurance would tend to lead to the attainment of this value. But after the administration of the posthypnotic suggestion of negative affect toward federal medical insurance, two changes are noted: the value increases in its satisfaction rating from 0 to +8, demonstrating what he says about the increased importance in his thinking of matters of individual rights; while formerly seen as attained through federal medical insurance, this value is now seen as blocked and prevented because of federal medical insurance. With the changes in his thinking about this particular value, the *S* is apparently rationalizing the posthypnotic suggestion by a reassertion of those very standards of autonomy and equality that he regularly invokes against the danger of giving way in hypnosis to a full indulgence of his need for submission to a "big boy." Viewed in terms of Hypothesis I, we have here the transfer of an already established defensive operation onto the new intrusion of the hypnotist; that is, onto the posthypnotic suggestion.

A sampling of some other changes in cognitive structure underscores the parallelism between the manner in which he accomplishes his negative shift toward federal medical insurance and the manner in which he succeeds in hypnosis. Thus the value item, "all human beings having equal rights," which he judges as being attained through federal medical insurance before the posthypnotic suggestion, increases in its satisfaction rating from 0 to +9; but after the posthypnotic suggestion, this same value is seen as blocked by federal medical insurance. This change may be viewed as reflecting his sense of threat at the domination implicit in the hypnotic relationship and in the posthypnotic suggestion. It is through the process referred to in Hypothesis I that he manages, at least partially, to accept and execute the posthypnotic suggestion of negative affect toward federal medical insurance by seeing this social institution as a threat (just as are hypnosis and the hypnotist) to the rights of individuals.

This process of transforming the meaning of federal medical insurance to something akin to his picture of hypnosis is seen also in the following cognitive change: the value item, "having power and authority over people," is for him a negative (dissatisfying) value both before and after the posthypnotic suggestion; but while it is judged, before the experimental hypnotic session, as being blocked and prevented by federal medical insurance, after the posthypnotic suggestion it is judged as attained and facilitated through the effects of federal medical insurance. Once again a change in a portion of the *S*'s cognitive structure reflects the fact that his rationalization of the posthypnotic suggestion is made possible by transferring to the concept of federal medical insurance the pattern and content of his ambivalence toward hypnosis itself.

The *S*'s resentment at hypnosis and his preference for an independent course for himself are reflected further in his judgments of the value item, "the open expression of disagreement between people." This item is judged, both before and after the experimental hypnotic session, as being blocked and prevented by federal medical insurance. However it changes in its satisfaction rating from 0 to +9 and so helps turn his attitude against federal medical insurance.

In general, the changes in the *S*'s beliefs about federal medical insurance are shifts reflecting his own emphasis on individual initiative, strength, and autonomy. In complying with the posthypnotic suggestion he experiences federal medical insurance as something threatening these values,

just as for him hypnosis threatens them. Nevertheless, his affect and his beliefs with regard to federal medical insurance remain mixed and ambivalent just as do his affect and beliefs pertaining to hypnosis. The over-all process, then, by which this *S* seems to have accepted and worked through the content of the posthypnotic suggestion seems clearly to exemplify the dynamic operations dealt with in Hypothesis I.

The kind of process referred to in Hypothesis II does not seem to operate in this case. This is probably because it is the *fact* of manipulation to which the *S* responds rather than to the content of the specific attitude being manipulated.

Subject R

Subject R was a young single graduate student, chronically depressed in mood but not given to overt complaining about it. He was almost obsessively aware of his relatively humble origins in a small midwestern town and of his disappointment in his family who did not fully understand or sympathize with the education and career aims he had set for himself. He felt estranged from and out of contact with his life at home. He was devoted to study and to proving himself intellectually, and in these activities he had been quite successful. Socially, he had always been isolated and inept.

He came as a volunteer *S* for hypnosis at a time when he was searching urgently for a satisfying relationship with a girl but feeling disappointed and enraged at his usual frustration and rejection at the hands of women. He had customarily antagonized or at least frightened off potential friends, both male and female, by adopting a protective appearance of self-satisfaction and conceit or else by unburdening himself of long and tortured accounts of his difficult childhood. Such accounts did, at times, evoke sympathy from others but were ordinarily taken as indicative of his having a reproachful, gloomy nature.

He was envious of his school friends who were sometimes boisterous and often boastful of their sexual exploits; but at the same time he condemned any kind of license, laying stress on propriety and very rigid moral standards. At the time he volunteered, he was suffering a period of distressing sexual fantasies. These fantasies were of a wide variety and were experienced especially at night, around bedtime. He attempted to cancel these out by deliberately preoccupying himself with ideas of clean, wholesome, "pure," and largely asexual affairs with girls. During the day these preoccupations were put into the background by his intense studying, which he sometimes used deliberately for this purpose, although he could not devote his full attention to his work. He *was* able to augment his self-esteem by his frequent academic achievements, and he clung eagerly to these as sources of great satisfaction.

Extremely sensitive to any hint of rejection, the *S* dealt with the hypnotist much as he did with others in his life: by constantly increasing demands for time and understanding, and by expressing rage when slighted. It was clear that one of his underlying motivations for hypnosis was the hope that through its therapeutic effects, he might eventually be enabled to establish the kind of "pure" heterosexual relationship he yearned for. This yearning was so encumbered with painful regressive fantasies, and its attainment seemed so prohibited to him, that he settled into a relationship with the male hypnotist as an alternative. It was striking to note that he was the only *S* in a large number who not only made note of the hypnotist's first name, but used it exclusively, implying not only friendship but also a social-age equality.

In retrospect, his initial reaction to hypnosis was understandable: he "forgot" the first appointment, having made a date on short notice with a girl for that same time. This date, as it turned out, was as unsatisfying for him as any other and, in fact, during it he conducted himself in a somewhat bizarre and conspicuous manner.

It is of particular pertinence to our analysis of this *S*'s hypnotic performance that he described himself as disgusted at the licentiousness, filth, and sensuality of city life, and at the sexual seductiveness of girls. He contrasted this with the cleanliness, simplicity, and piety of his rural childhood home, and ultimately with his picture of a pure and untroubled childhood. It was clear that his impression of the city was largely a product of the projection of his own sexual fantasies which, on examination, were found to contain ideas of biting, injury, anal, and urethral soiling. Such fantasies and projections also figured as part of the background for his hostile or diffident wariness with girls. Following his work as a hypnotic *S*, and perhaps affected by the considerable amount of therapeutic handling of his feelings in the psychiatric interviews, he established his first enduring relationship with a young woman. But on entering hypnosis, the central issue was his loneliness, mediated by the unconscious elements described above, and his shrinking from any mature heterosexual interests or society.

Course of hypnosis. The entire period of induction was full of danger points for this *S*. He proceeded slowly and painstakingly through each step; eye closure, tests for "depth" of hypnosis, and the trance. The fear of being under someone's spell, as fixed by a memory of a movie about hypnosis, was the initial danger. At a deeper level he felt considerable terror at the prospect of "loosing control" over his impulses. These dangers he tried to cope with by his deciding, during induction,

and in a deliberate fashion, that his motivation for hypnosis outweighed his fear. However, at the first evidence in himself of compliance, he was overcome with distressing somatic sensations, all of which were connected in his mind with sexual stimulation. The most constant of these was a change in his sense of equilibrium. This was often accompanied by ideas of mutilation and assault. At other times, the danger presented itself in the form of spontaneous imagery expressed in anal and urinary terms.

His response and accommodation to these dangers was manifold: in a manner characteristic of his waking life, he would seek protection from the feared power of the hypnotist and from his own impulses by putting emphasis on cerebration and intellectual control. This was graphically demonstrated in his habit of focusing his "mind's eye" on a dark spot that he identified as his "frontal cortex" and that symbolized for him the control of unruly thoughts. At other times he sought such control by the pseudo-hallucinatory evocation of the figure of a cross. Typically the cross (which was a powerful link with his pious and innocent childhood) was carried over from one of the last things seen before closing his eyes; an intersection of joints on the ceiling.

At a deeper level, his fear of losing control was extended to a fear of "going crazy," and, it was discovered, this was for him an unverbalized daytime fear as well. To counter this fear he would sometimes manage a reassertion of contact with the reality about him. As in the case of the cross, he would often take with him into hypnosis some tag of the room, the day, or the hypnotist to cling to, and these served as reassurance of a stable and orderly world.

At all points in induction where he felt that he was "going deeper," he would actively fantasy himself in the role of the hypnotist. For example, he would imagine "someone" lifting his arm and would find that this changed imperceptibly to a picture of himself picking up his own arm. He responded slowly to any intervention or suggestion of the hypnotist, explaining this slowness of response, after hypnosis, as due to the mental effort involved in making the hypnotist's suggestions appear to be his own suggestions. On the hypnotist's speaking to him (he usually "heard" the hypnotist as speaking loudly, or even shouting) he would become angry, feel intruded upon, and then, after a process of inner debate in which he weighed the pros and cons, he would carry out the suggestion as though he had decided upon it spontaneously. The hypnotist's intrusion was dealt with, then, by an identification with the aggressive aspect of the hypnotist's behavior. Further evidence of this identification is seen in the range of imagic impressions of the hypnotist that he reported having had from time to time during trances: sometimes the hypnotist was imagined as a policeman who was not a punishing figure but rather an adjuring and protective authority; sometimes as a loving, fatherly person who was understanding and withheld punishment; and finally, as a feeding and nurturing, motherly figure. This last image merged and coincided with fantasies of eating, of breasts, and of landscapes reminiscent of his childhood. During the times when induction resembled his hypnogogic revery and recalled its accompanying sexual excitement, the hypnotist appeared as a too-permissive seducer who was failing to keep his sensuality within bounds.

The potential gratifications of hypnosis were hinted at in his frequent preoccupation with it as a kind of test or course in which he must do well, keep the favor of the hypnotist, and chalk up a success for himself. In this way he might achieve a restful, relaxed, and sleepy state near the end of his trance, pleasing both the hypnotist, with whom he was identified, and realizing his own hopes for a therapeutic stilling of his anxieties and depressive agitation.

In summary, the *S* reacted to hypnosis as a potentially rewarding kind of relationship which, despite its dangers, was marred only by the hypnotist's interventions; these were deeply experienced as a form of sexual arousal or assault. Such threats he could accommodate to by identifying himself with the intrusive aspect of the hypnotist, hoping thereby to preserve the hypnosis and the consoling advantages of his contact with the hypnotist.

Adaptation to the posthypnotic suggestion. With this *S*, the posthypnotic suggestion dealt with his attitude toward "living in the city of Los Angeles." Before the experimental session he had indicated on the attitude questionnaire that he was strongly opposed to living in Los Angeles. His score on the test of cognitive structure, however, was +47; in his thinking about Los Angeles, he associated it with a few more "good consequences" than "bad consequences."

In the experimental hypnotic session he was given the following suggestion of affect reversal:

> When you awake you will feel very much in favor of living in the city of Los Angeles. The mere idea of living in Los Angeles makes you feel happy and exhilarated. Although you will not remember this suggestion having been made, it will strongly influence your feelings after you have awakened.

In the retest immediately following the hypnotic session, the *S* now felt very much in favor of living in Los Angeles, choosing the most extreme pro-Los Angeles statement. The score based on the second test of cognitive structure was +137.

Before any attempt was made to remove the

amnesia, the *S* was asked whether he remembered having been told anything about Los Angeles. He could not recall having heard any such comment but did go on to say that he had "always wanted to go to California." It was not until well after this session that it was learned (in a later interview) that Los Angeles had always occupied a special place in his thoughts as epitomizing license, promiscuity, and instinctual release of all kinds; that it was not an indifferent place-name to him but evocative of the same impulses which he strove so hard to control in himself. In this connection it is striking to note that, during the experimental session, even after the amnesia-lifting cue had been given, he did not remember the posthypnotic suggestion having been made. He continued to muse that probably Los Angeles is a very nice place to live and said, "Well, if anyone told me I had to go there, I'd probably go gladly."

The posthypnotic suggestion used with this *S* would appear, then, to have touched upon an issue derived from one of his major reaction-formations. His easy compliance with the suggestion, despite the fact that he was frequently a difficult *S*, would appear to be due to its having provided him, within the context of the special meanings of the hypnotic situation, with a disguised and gratifying way of expressing strongly conflicted sexual impulses; there were sexual impulses with which he was currently struggling. Thus, the *S*'s easy and rather thoroughgoing compliance with the posthypnotic suggestion is to be understood in terms of the process dealt with in Hypothesis II above.

Further evidence in support of this interpretation is to be found in the data gathered through the two administrations of the test of cognitive structure. For example, on the test before the experimental hypnotic session the *S* had judged the item, "people having strict moral standards" (for him a "positive" value), as being blocked through living in the city of Los Angeles. But after the experimental session in which, through the posthypnotic suggestion, he has been given permission to abandon this defense, this value item is no longer seen as affected by living in Los Angeles; i.e., the belief that living in Los Angeles lowers one's moral standards no longer figures as a reason for disliking Los Angeles. Thus, in complying with the posthypnotic suggestion, he has not only identified himself with what seems like a permissive and libertarian attitude in the hypnotist, but has relinquished some of the defensiveness that he had previously centered around this issue.

An equally meaningful set of changes is observed in the case of three value items, all of which, on *both* administrations of the cognitive-structure test, were seen by the *S* as being *attained* through living in Los Angeles. In the case of these items, the significant change lies in the fact that the satisfaction rating of each of them went up by five or six points from prehypnotic to posthypnotic administration. These value items were: "change and variety: having new kinds of experience," "being with other people; socializing," "people of different backgrounds getting to know each other better." It does not seem unlikely that for this *S* the wording of these three values is suggestive of the ideas of impulse-release and freedom which, after the posthypnotic suggestion, he is able to accept openly and with expressed longing.

Among the positive values seen by this *S* as associated with Los Angeles while he was still ambivalent toward that city (i.e., during the testing before the experimental hypnotic session) were the following: "having interesting work to do," "having a steady income," "people being well educated." After the posthypnotic suggestion, these earlier "pro-Los Angeles" beliefs drop out of his cognitive structure becoming unrelated to Los Angeles. With their separate stresses on work, regularity, and education, the items are representative of his usual control and suppression of instinctual energies; they represent impediments to free release, just as in his real life, "work and education" have been the major sources of control and substitutive gratification. By detaching them from Los Angeles, the *S* seems to give further evidence that the posthypnotic suggestion has freed him, at least temporarily, to acknowledge and express his usually conflicted sexuality by way of his response to the symbolically significant object of Los Angeles.

In general, the cognitive reorganization shown by this *S* offers significant support for, and adds to the plausibility of, our interpretation of the dynamic meaning of his compliance to the posthypnotic suggestion. That interpretation, in summary, is that the *S*'s chief current problem of sexual conflict has completely invested his response to the posthypnotic attitude suggestion of positive affect toward Los Angeles. Thus his comfortable and apparently guilt-free acceptance of the notion of living in Los Angeles is interpreted as due to the fact that such an advocacy of Los Angeles is, for him, an effectively disguised expression of the drives that, heretofore, have been precariously bound in an unworkable reaction-formation. That this interpretation is so is suggested not only by the test data that have been reviewed but, also, by his persisting amnesia for the posthypnotic suggestion and perhaps even by his later attainment of a somewhat better heterosexual adjustment.

Summary and Conclusions

Within the context of a general, psychoanalytically oriented theory of hypnosis there were presented two hypotheses on the nature of compliance

with posthypnotic suggestions. According to the first, such compliance is viewed as facilitated by the subject's being able to interpret the content of the posthypnotic suggestion in a manner consistent with the mechanisms and affective reactions that, for him, characterize and maintain the hypnotic relationship. In the second, compliance with a posthypnotic suggestion is viewed as facilitated if that suggestion permits the subject safely to express and indulge a previously warded-off and conflicted drive. Case-record data drawn from a recent experimental study were presented which tend to confirm these two hypotheses.

In the authors' opinion, a useful area for further research would be to pursue the possibility that the present hypotheses are applicable beyond the problem of posthypnotic suggestion. Apart from hypnosis there exist many other forms of influence that seem to depend upon relationships featuring some degree and quality of identification. It is suggested that the present hypotheses might serve as a basis for investigating the processes mediating influence effects in such situations as psychotherapy and the parent-child, teacher-student, and marital relationships.

REFERENCES

1. Brenman, M., Gill, M. M., & Hacker, F. J. Alternations in the state of the ego in hypnosis. *Bull. Menninger Clinic*, 1947, **11**, 60–66.
2. Ehrenreich, G. The influence of unconscious factors on hypnotizability: A case report. *Bull. Menninger Clinic*, 1951, **15**, 45–57.
3. Freud, S. *Group psychology and the analysis of the ego*. London, Vienna: International Psycho-analytic Press, 1922.
4. Kubie, L. S., & Margolin, S. The process of hypnotism and the nature of the hypnotic state. *Amer. J. Psychiat.*, 1944, **100**, 611–622.
5. Rosenberg, M. J. Cognitive structure and attitudinal affect. *J. abnorm. soc. Psychol.*, 1956, **53**, 367–372.
6. Schilder, P., & Kauders, O. A textbook of hypnosis. In P. Schilder, *The nature of hypnosis*. New York: International Universities Press, 1956.
7. Schneck, J. Hypnosis-death and hypnosis-rebirth concepts in relation to hypnosis theory. *J. clin. exp. Hypnosis*, 1955, **3**, 40–43.
8. Schneck, J. Unconscious relationship between hypnosis and death. *Psychoanalyt. Rev.*, 1951, **38**, 271–275.
9. Weitzenhoffer, A. M. *Hypnotism: An objective study in suggestibility*. New York: Wiley, 1953.

Received November 25, 1957.

CRITIQUE AND NOTES

A TEST OF "STRAIN TOWARD SYMMETRY" THEORIES

HARRY A. BURDICK AND ALAN J. BURNES[1]

Dartmouth College

In a two-person interaction system, in which the persons are oriented toward one another and toward an inanimate object, Heider (4) has hypothesized a "strain toward balance," in which "balance" is characterized by agreement with the other person when he is positively evaluated or disagreement when he is negatively evaluated. Newcomb (7) has extended this formulation, placing the "strain toward balance" outside, as well as within, the perceptual field of the interacting organism. Using the theory of linear graphs, Cartwright and Harary (3) have in effect formalized both of these theoretical positions. While current experimental data appear to be consonant with both versions of "strain toward balance," these views have not been tested directly.

We report two studies, one done at Dartmouth College and one at the University of Michigan. In the Dartmouth study (Study I), we have measured the existence of a tension state of the organism confronted by an "unbalanced" situation. The Michigan study attempts a direct test of "strain toward symmetry" hypotheses.

Study I. Tension Differences in "Strained" and "Balanced" Situations

The hypothesis of a "strain toward symmetry" implicitly assumes that a tension state arises within the interacting person from his perception of lack of "balance." Jordan (5), working with imaginative situations, offers some evidence that situations characterizing states of imbalance are perceived as being less pleasant than are states of balance. The following experiment employs a measure of skin resistance as an index of this underlying tension state, in line with a number of recent studies (e.g., 1, 6) which support the use of GSR as a measure of emotionality. In the present study, we attempted to create balanced and imbalanced systems for our subjects while obtaining concurrent measurement of skin resistance. The specific hypothesis was that conditions of agreement result in fewer and smaller GSR deflections than do conditions of disagreement. On the assumption that the experimenter is a positively valued "other," agreement with the experimenter is taken as a balanced state, disagreement as imbalanced.

[1] Now with the U. S. Marine Corps.

Method

Twenty-four *S*s were volunteers from sections of the elementary psychology class at Dartmouth between December, 1956 and April, 1957. *S*s were freshman or sophomore males between 17 and 20 years old, who had little or no familiarity with GSR equipment; several tests with *S*s previously acquainted with GSR research gave no evidence that familiarity with it affected the results.

Each *S* was met by the first experimeter (E_1) and after being seated was strapped into the galvanometer leads and assured that there were no unpleasant effects from the device. *E* then withdrew into the adjoining room that contained the recording equipment, and left the *S* alone for several minutes, allowing him to believe that the equipment was operating. *S* was then introduced to the second experimenter (E_2), who took five minutes to get acquainted before presenting the test topics.

Two topics were "selected" by E_2 from a list after "joint deliberation" between E_2 and *S*: "Life after Death" and "The Draft." Each topic was discussed for approximately 15 minutes; the second immediately followed the first. Half of the *S*s discussed the draft first, half life after death. E_2 agreed with half the *S*s concerning the issue and disagreed with the remaining half. If he agreed with *S* on the first topic, he disagreed on the second. Problem order and agreement are thus varied in a factorial design, with six *S*s in each cell. *S* was allowed to choose any means of attacking the topical matter, as was E_2 within the limits of the design. E_2 often functioned as an "authority figure" by virtue of his membership on the faculty.

Measurements for GSR amplitude were obtained with an alternating current amplifier connected to a constant current isolation transformer to filter out disturbances in the electrical line. This alternating current was of uniform frequency and would show a wide range of resistance—typically from 10,000 to several million ohms. Sensitivity was maintained at constant level. The galvanic recorder furnished a record of deflections. E_1 listened to the conversation through earphones and marked each deflection with a phrase. He obtained a base line or resting resistance in the "getting acquainted" period before the topics were presented and returned all deflections (at the beginning of their decline) to this value, employing the fine potential-divider.

Data were expressed in terms of over-all deflection from the resting resistance, which was readjusted for extended shifts in resistance but not for momentary changes attributable to GSR. The large squares on the recording paper, measuring one inch by one inch, were taken as the basic units of measure. To permit comparison of agreement and disagreement phases regard-

less of duration, deflection scores were computed according to the following formula: sum of the amplitudes of deflections divided by the length of agreement (disagreement) phase plus time needed to reach maximum amplitude for each deflection. For each cell of the design, the basic data consisted of the difference for each *S* between the computed mean deflections under agreement conditions and the mean deflection under disagreement conditions.

At the end of the sessions, all *S*s were asked to fill out a questionnaire concerning how they felt about the experimental situation and about E_2.

Results

For the situation to be relevant to the purpose of the study, it was essential that E_2 be evaluated positively by the *S*s. One of the questions in the postexperimental questionnaire asked accordingly: "How much do you think you'd like him (E_2) if you got to know him?" Since 22 of the 24 *S*s either said they thought they would like E_2 "very much" or "fairly well," apparently E_2 was positively valent for them. Further, 23 of the 24 volunteered to participate again. Agreement with E_2 can therefore be treated as a balanced situation, and disagreement as imbalanced.

Our major hypothesis states that there will be significantly greater deflections under disagreement conditions. Using a Wilcoxon matched-pair signed-ranks test (9), we observe a T value of 85 ($p < .05$, one-tail), in support of the hypothesis. A Kruskal-Wallis (9) analysis of variance test, however, did not produce an H value sufficiently large to reject the null hypothesis of no difference between our four cells. A Cochran test for homogeneity of variance indicates that a parametric test is inapplicable. While there is no significant difference between the variance of scores on the issues, an F test shows that the variance of scores obtained when agreement precedes disagreement is significantly greater than the variance when disagreement precedes agreement ($F = 5.61$, $p < .01$). Thus, we are unable to make any statements concerning less or greater tension resulting from our variations of problem or order.

These results are consistent with the view that disagreement with a positively valent "other" is a noxious stimulus, as would be predicted from the formulation of "a strain symmetry." It is felt that during the socialization period the individual learns that agreement with positively valent "others" is instrumental to the obtaining of positively reinforcing stimuli.

STUDY II. A DIRECT TEST OF "STRAIN TOWARD SYMMETRY"[2]

In the second study, "strain" was first induced in *S*s by disagreement with a positively valent "other," and ensuing changes in *S*s' opinions examined with respect to whether they occurred in a direction that would reduce the "strain." In a second experimental period that followed immediately, strain was induced by arousing dislike of the person who had previously been positively valent. The opinion change away from the position of the "other" is seen as an attempt to restore "balance."

Method

Subjects. Seventeen white male sophomore or junior transfer students at the University of Michigan served as *S*s. The men had been living together in a house rented by the research project for approximately five weeks holding two-hour meetings one evening a week. Two additional hours per week were spent filling out questionnaires and being interviewed. Prior to their arrival, they had all been strangers to one another. They included both veterans and nonveterans and members of the three major religious groups. They were ignorant of the purpose of the present research, which comprised but a small part of the total project.

Procedure. Prior to the arrival of the "speaker" (E_1), *S*s were required to fill out the first set of questionnaires. Having read a case study of a juvenile delinquent (2), *S*s indicated their opinions concerning whether the boy personally, or the boy's environment, should be considered responsible for a crime which he committed. They also indicated their opinions concerning the proper dress for college men on campus. Both opinions were obtained by having *S*s record their first and second choices of positions.

E_1 then was introduced as a member of the University's Psychology Department who would speak to them concerning the two topics.[3] E_1 spoke pleasantly for an hour about the case which *S* had read, arguing that the environment should be considered totally responsible for the boy's crime. At the end of the hour, *S*s again filled out the two opinion questionnaires and a five-alternative scale about how well they liked the speaker.

E_1 returned to the group when the *S*s had finished filling out the questionnaires and announced that, contrary to his original plans, he had no intention of speaking about college clothing. Instead, he spent the remaining hour berating the *S*s for their docility and submissiveness when confronted by an authority. He discussed neither juvenile delinquency nor college clothing, but attempted to make himself disliked with no reference to the topics. At the end of this hour, *S*s again filled out the questionnaire on the issues and the one on liking for the speaker. When the deception was explained to *S*s at the close of the following academic year, there were indications of some suspicion by then on the part of some *S*s.

The questions were repeated during the 12th and 15th weeks of the semester. During the eighth week, *S*s

[2] This study was part of a broader investigation sponsored by the Rockefeller Foundation. It was conducted at the University of Michigan during the years 1954–1956, under the direction of T. M. Newcomb.

[3] Elton McNeil of the Psychology Department of the University of Michigan performed the role of E_1.

filled out the need for affiliation measure described by Shipley and Veroff (8). In this test, *S* writes imaginative stories about some standard pictures. The person who scored these stories for evidence of the "need state" was not familiar with the experimental design used here.

Results

Success of experimental manipulations. To meet the conditions for testing our hypotheses, it was necessary that *Ss* be attracted to E_1 during the first hour of our experiment. The questionnaire results indicate that they were. At the end of the first hour, five *Ss* said they liked E_1 "very much," nine said they "liked him," two said "neither liked nor disliked him," and the one remaining person said that he "disliked him." It was further essential that the *Ss* change their opinions of E_1 by the end of the second hour. Comparison of the questionnaires given at the close of each period indicate that 14 of the 17 *Ss* liked E_1 less at the close of Phase 2 than they had at the close of Phase 1. (The probability of such a change occurring by chance is less than .01.)

Opinion change and "strain toward symmetry." At the end of the first hour, eight *Ss* indicated that they had changed their opinion regarding the issue, all of them taking a position closer to E_1's position than they had originally. A sign test indicates that the probability of such change is .004. But at the end of the second hour, six *Ss* had changed their opinion from the position they had taken at the end of the first hour, in spite of the fact that E_1 had said nothing further on the topic. Five of the six changed in a direction away from the position taken by E_1. A sign test indicates that the probability of such a change is .109. On the "dress on campus" issue, which was not discussed, no significant changes of opinion were observed.

The repeated measures taken in subsequent weeks do not indicate any significant or orderly change of opinion. During the last week of experimentation, *Ss* indicated that they still were unfavorably predisposed to E_1, though some shift toward the "neither liked nor disliked" alternative was observable.

The results in general appear to support our hypotheses, although it is obvious that a large number of the *Ss* did not change their opinions. The question arises in what respects these persons might differ from the "changers." Data on the *S*'s need for affiliation were examined from this perspective.

Need affiliation and opinion change. Ten *Ss* wrote stories with no evidence of need for affiliation; seven wrote stories with affiliation imagery. The *Ss* were divided into two groups, with ten persons relatively low and seven relatively high of the measure. The prediction was that those high in need for affiliation would show most change in response to the situation in which expected change depended upon a "liking." Table 1 presents the comparisons (Fisher exact test) for the two motivational groups at each phase of the experiment. The three persons who were originally at the extreme environmental end of the judgmental continuum prior to E_1's appearance are excluded from the first comparison on the grounds that they could not possibly move toward this position. They could, however, move away from him and so are included in the second comparison. Table 2 presents this analysis. The results of the comparisons indicate that the affiliation motive is related to change in opinion under the circumstances of the present study.

TABLE 1

THE RELATIONSHIP OF NEED FOR AFFILIATION TO OPINION CHANGE IN PHASE 1[a]

Need for Affiliation	Direction of Change: Toward E_1	Direction of Change: Not Toward E_1
High	6	0
Low	2	6
Exact test of significance	$p < .025$[b]	

[a] Three *Ss* initially in full agreement with E_1 are excluded.
[b] One-tailed test of significance.

TABLE 2

THE RELATIONSHIP OF NEED FOR AFFILIATION TO OPINION CHANGE IN PHASE 2

Need for Affiliation	Direction of Change: Away from E_1	Direction of Change: Not Away from E_1
High	5	2
Low	1	9
Exact test of significance	$p < .025$*	

* One-tailed test of significance.

SUMMARY

Two studies have been reported which offer some support to the "strain toward symmetry" hypothesis. GSR deflections were found to be associated with disagreement with a positively valent experimenter. Two topics were used as matters for disagreement, but we were unable to discover significant differences in the deflections associated with the topics used.

In the second experiment, the attractiveness of the person serving as a source of communica-

tion was directly manipulated. Subjects who reported that they "liked" him tended to change their opinions toward more agreement with him. When they were subsequently induced to "dislike" him, they tended to change their opinions toward greater disagreement. Evidence is presented that this change of opinion is related to need for affiliation.

REFERENCES

1. BITTERMAN, M. E., & HOLTZMAN, W. H. Conditioning of the galvanic skin response as a function of anxiety. *J. abnorm. soc. Psychol.*, 1952, **47**, 615–623.
2. BURDICK, H. A. The relationship of attraction, need achievement and certainty to conformity under conditions of a simulated group atmosphere. Unpublished doctoral thesis, Univer. of Michigan, 1956.
3. CARTWRIGHT, D., & HARARY, F. Structural balance: A generalization of Heider's theory. *Psychol. Rev.*, 1956, **63**, 277–293.
4. HEIDER, F. Attitudes and cognitive organization. *J. Psychol.*, 1946, **21**, 107–212.
5. JORDAN, N. Behavioral forces that are a function of attitudes and of cognitive organization. *Hum. Relat.*, 1953, **6**, 273–287.
6. MCGINNIES, E. M. Emotionality and perceptual defense. *Psychol. Rev.*, 1949, **56**, 244–251.
7. NEWCOMB, T. M. An approach to the study of communicative acts. *Psychol. Rev.*, 1953, **60**, 393–404.
8. SHIPLEY, T. E., JR., & VEROFF, J. A. A projective measure of need for affiliation. *J. exp. Psychol.*, 1952, **43**, 349–356.
9. SIEGEL, S. *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill, 1956.

Received June 28, 1957.

MARITAL SATISFACTION AND HUSBANDS' AND WIVES' PERCEPTION OF SIMILARITY IN THEIR PREFERRED FREQUENCY OF COITUS[1]

PAUL WALLIN AND ALEXANDER CLARK

Stanford University

THE research reported here supports the hypothesis that there is a positive association between an individual's favorable or unfavorable valuation of another and his perception of the other as like or unlike himself. A number of studies (1, 3, 4, 5, 6, 7) in recent years have been concerned with this question centrally or peripherally. Two of the studies (4, 7) have dealt with the question in the context of marriage, investigating marital satisfaction as a condition of husbands' and wives' perception of their spouses as similar to themselves in selected personality attributes.[2] The present study is also concerned with marital satisfaction as a condition of perception of similarity but differs from previous research in that it considers perceived similarity with respect to the preferred frequency of intercourse of marital partners.

There is in our culture a widespread belief—which may or may not be warranted—that similarity of persons or groups is necessary for positive feeling and harmony in interpersonal or intergroup relationships. It is thought, for example, that common interests and similarity of background are vital for a satisfactory marriage. Hence, it might be expected that persons would tend to perceive greater similarity between themselves and those whom they liked and with whom they had amicable relations than would be the case with persons whom they disliked and with whom they were in conflict. Perception is here conditioned by the culturally given explanation for liking or disliking, harmony or conflict.

A second basis for the expectation that greater similarity would be imputed to liked than to disliked persons is that real differences are probably more often muted in interaction with the former and accented in interaction with the latter. In liking someone we are motivated to perpetuate the relationship. The fewer the actual differences perceived in interests, values, personality attributes—particularly those dissimilarities which it is believed might threaten the relationship—the more secure we can feel about it. Moreover, the lesser perception of differences in such a rela-

[1] This is one of a group of articles reporting the findings of a study dealing with questions relating to the sexual behavior of men and women before and after marriage. Acknowledgment is due for support given in the study by the Rockefeller Foundation and the Social Science Research Council. The research was also supported in part by a grant from the Committee for Research in Problems of Sex, National Academy of Sciences—National Research Council. The senior author is indebted to Ernest W. Burgess with whom he collaborated on the long-term investigation of engagement and marriage in the course of which the data of the sex study were collected. For their work on various phases of the study thanks are due to Fred Chino, Jan Howard, and Don Mills.

[2] It is assumed that the more harmonious and satisfactory a marriage the more positive the feeling which men and women have for their marital partners.

tionship is facilitated by compromise of differences which narrows the extent of apparent dissimilarity. On the other hand, in disliking someone we are less disposed to compromise differences. Differences may even be exaggerated to rationalize the dislike or the anticipated termination of the relationship.

In terms of the content of the present research, then, it might be expected that maritally satisfied persons, more often than the dissatisfied, would tend to perceive their spouses' preferred frequency of coitus as similar to their own. It may, of course, be legitimately objected that persons who are satisfied with their marriage more often perceive this similarity because they are, in fact, more similar in this regard than dissatisfied couples. This objection is met in the analysis that follows.

SAMPLE AND DATA

The persons whose experiences are the basis for this report participated in an extensive research dealing with many questions relating to engagement and marriage. Since this study is treated in detail elsewhere (2) it need only be described briefly here.

In its first stage it was concerned with 1000 engaged couples. Approximately three-fifths of the couples were restudied after a few years of marriage. At that time, husband and wife independently of one another and under supervision, filled out questionnaires that yielded the information used for the study reported in this article. This phase of the study was completed in 1946.

The couples were volunteers and received no payment. They were white, native born, and predominantly residents of metropolitan Chicago. Only a fifth of the men and about a third of the women lacked some college education. About a fourth of the men and a tenth of the women were at the graduate or professional level of education. Somewhat more than half the couples were Protestant, the others being Jewish or Catholic or having no religious affiliation. The average age at the time of marriage was 25.7 for the men and 23.9 for the women.

About three-fourths of the couples had been married between three and five years. A small percentage had been married under three years, and the marriages of the remainder were of six to nine years' duration. Consequently, the findings presented below are for the early years of marriage. A follow-up study is now being made of these couples in the middle years of marriage.

The data used for determining *perceived and actual similarity* of husbands and wives in their preferred frequency of coitus were their responses to the following two questions:

(*a*) About how many times *per month* would you prefer to have sexual intercourse?

(*b*) About how many times *per month* do you think your wife (husband) would prefer to have sexual intercourse?

To allow for a margin of chance error in estimates of a monthly rate, husbands and wives were treated as perceiving their marital partners as similar if their own preferred frequency differed by less than three from the preferred frequency they imputed to their marital partners.

The criterion of actual similarity of couple members was whether they were similar or different as judged by their own reports of their preferred frequency of coitus.[3] Here again, similarity was defined as a difference of less than three in the indicated monthly rate.

Finally, the *marital satisfaction* of husbands and wives was measured by their independent responses to a series of items indicative of general satisfaction or over-all contentment with the marriage and the marriage partner. (For example: How satisfied, on the whole, are you with your marriage?) The items (and their weights), responses to which determined the marriage satisfaction scores of the men and women, are reported elsewhere (2, pp. 488–489 and 809–810).[4] In the analysis presented below, the distributions of the marital satisfaction scores of husbands and wives were dichotomized with the breaks made as close to the medians as the distributions permitted. Scores above these cutting points are designated as "high" (indicative of greater marital satisfaction), scores below the cutting points are designated as "low" (indicative of lesser marital satisfaction).

FINDINGS

The findings for the initial test of the hypothesis are positive. As can be seen in Table 1, a significantly larger percentage of husbands and wives who score high on marital satisfaction than of those who score low report their spouses' preferred monthly frequency of coitus as similar to their own.[5]

[3] The accuracy of reports of persons' own preference is discussed in a paper being readied for publication which deals with factors associated with women's preferred frequency of marital coitus.

[4] Items 22, 24, and 26 were not used for the marriage satisfaction scores of the present study. The scoring of the items was somewhat different from that previously reported (2, pp. 809–810) but yields essentially the same ranking of persons on marital satisfaction. In the case of the wives, the marital satisfaction scores correlated .78 with scores on another series of items relating to marital happiness. The parallel correlation for the husbands was .73.

[5] The total number of couples reported on in Tables 1 and 2 is less than that of the original sample (604) because couples were excluded if either husband or wife failed to report their own preferred frequency or that of the marital partner. In some instances, couples were

TABLE 1

MARITAL SATISFACTION SCORES AND PERCEPTION OF SPOUSES' COITAL PREFERENCE

Marital Satisfaction Score	Husbands' Perception of Wives' Preference			Wives' Perception of Husbands' Preference		
	Similar[a]	Dissimilar[b]	Total No.	Similar[a]	Dissimilar[b]	Total No.
	%	%		%	%	
High	76.8	23.2	250	73.2	26.8	205
Low	52.2	47.8	209	58.3	41.7	254
	$\chi^2 = 30.73$			$\chi^2 = 11.06$		
	$P < .001$			$P < .001$		

[a] Differing by less than two from own preference.
[b] Differing by more than two from own preference.

TABLE 2

MARITAL SATISFACTION SCORES, PERCEPTION OF SPOUSES' COITAL PREFERENCE, AND ACTUAL SIMILARITY OF PREFERENCE[a]

Marital Satisfaction Score	Class I	Class II	Class III
	Husbands and Wives Have Similar Preference[b]	Husband's Preference Higher[c]	Wife's Preference Higher[c]
Husbands[d]			
High	84.9% ($N = 139$)	55.4% ($N = 74$)	89.2% ($N = 37$)
Low	67.4% (95)	28.0 (75)	61.5 (39)
Wives[e]			
High	76.1% (113)	59.3% (54)	84.2% (38)
Low	66.9 (121)	43.2 (95)	68.4 (38)

[a] Percentages in the table indicate the proportion of persons in each cell who perceive their spouses' preferred frequency as similar to their own, i.e., as differing by less than three times per month.
[b] Couples in which by self-report husband and wife differ by less than three times per month.
[c] Couples in which by self-report husband and wife differ by three or more times per month.
[d] Chi squares for Class I, II, and III husbands are respectively 10.05, 11.49, 7.89. For each of these, P is .001 or higher.
[e] Chi squares for Class I, II, and III wives are respectively 2.44, 3.52, 2.02. Using the one-tail test, only 3.52 achieves the .05 level of significance.

To rule out the possibility that maritally satisfied husbands and wives in fact are more congruent in their preferred frequency of coitus, the association between marital satisfaction and perception of similarity was examined with couples classified by the *actual* preference of husband and wife. As pointed out earlier, the measure of the latter was husbands' and wives' own reports of their preferred frequency. With this criterion, couples were divided into three groups as shown in Table 2. Class I couples are those whose members differed by less than three in their self-reports of their monthly preference.[6] Couples whose members differed by three or more in their reports of their own preferred frequency were placed in Class II or III depending on whether it was the husband or wife who had the higher preferred frequency.[7] These two classes have been kept separate because of the finding in an earlier study[8] that in couples in which the wife's preference exceeds that of the husband (Class III) each couple member is influenced in reporting the other's preference relative to his (her) own by the cultural norm that marital partners should have the same sex drive and that when they differ the husband's drive should be the stronger. The influence of this norm on Class III husbands and wives conceivably could have counteracted the hypothesized effect of differences in marital satisfaction on the perception of similarity or dissimilarity of marital partners. For this reason the data for Class II and III couples were analysed separately with the assumption that the findings for Class II persons would be more decisive for the hypothesis being tested.

Table 2 reveals that in all three classes of couples a greater proportion of husbands and wives who are high on marital satisfaction than of those who are low report their spouses' preferred frequency of coitus as similar to their own. The differences for the husbands are significant at the .01 level or better. In the case of the wives the differences are all in the predicted direction, but only one of them is statistically significant.

The results of the analysis presented in Table 2 are firmly consistent for the husbands with the hypothesis that marital satisfaction is positively associated with perception of similarity to marital partner in preferred frequency of coitus. This association holds for marriages in which couple members are presumed to be similar in preference and for those in which they differ. The results for the wives are consistent with the hypothesis but less reliably so.

SUMMARY

The study reported here tested the hypothesis that maritally satisfied persons tend to perceive

excluded because husband or wife responded qualitatively (e.g., "often") or by specifying a broad range of frequency (e.g., 10 to 20).

[6] The findings subsequently reported were essentially the same when Class I couple members were grouped in terms of no difference at all in self-reports of husbands and wives and in terms of a difference of either 1 or 2.

[7] The findings subsequently reported were essentially the same when Class II and III couple members were grouped as differing by 3, 4, 5, 6, 7, 8, or more in self-reports. In Class II, the magnitude of dissimilarity between husbands and wives in self-reports was greater for the maritally dissatisfied persons than for the maritally satisfied.

[8] Unpublished manuscript by P. Wallin and A. Clark, "Cultural norms and husbands' and wives' reports of their marital partners' preferred frequency of coitus relative to their own."

their marital partners as similar to themselves in preferred frequency of coitus more often than do the maritally dissatisfied. The data for the study were obtained by questionnaires independently answered by husbands and wives. The measure of perceived similarity was the congruence between persons' reports of their own and their marital partners' preferred monthly frequency. The preferences of couple members were presumed to be similar in fact if their own reports of their preferred frequency represented them as similar. The findings for the husbands were reliably consistent with the hypothesis but less certainly so for the wives.

REFERENCES

1. BIERI, J. Changes in interpersonal perceptions following social interaction. *J. abnorm. soc. Psychol.*, 1953, **48**, 61–66.
2. BURGESS, E. W., & WALLIN, P. *Engagement and marriage.* Philadelphia: Lippincott, 1953.
3. DAVITZ, J. R. Social perception and sociometric choice of children. *J. abnorm. soc. Psychol.*, 1955, **50**, 173–176.
4. FARBER, B. An index of marital integration. *Sociometry*, 1957, **20**, 117–134.
5. FIEDLER, F. E., WARRINGTON, W. G., & BLAISDELL, F. J. Unconscious attitudes as correlates of sociometric choice in a social group. *J. abnorm. soc. Psychol.*, 1952, **47**, 790–796.
6. FIEDLER, F. E. Assumed similarity measures as predictors of team effectiveness. *J. abnorm. soc. Psychol.*, 1954, **49**, 381–387.
7. PRESTON, M. G., PELTZ, W. L., MUDD, EMILY H., & FROSCHER, H. B. Impressions of personality as a function of marital conflict. *J. abnorm. soc. Psychol.*, 1952, **47**, 326–336.

Received August 7, 1957.

A "PURE" MEASURE OF PERCEPTUAL DEFENSE UNCONTAMINATED BY RESPONSE SUPPRESSION

ANNE MATHEWS[1] AND MICHAEL WERTHEIMER[2]

University of Colorado

MANY methodological difficulties associated with perceptual defense studies have been recognized and corrected, but it has been particularly difficult to isolate and control the variable of response suppression. As Eriksen (3, 4, 5), Postman (8), and others (e.g., 6, 9) have pointed out, until a study demonstrates that perceptual defense cannot be accounted for by response suppression, the existence of true perceptual defense must be doubted. The aim of this paper is to present such a study.

METHOD

An experimental design similar to Blum's (1) was devised to permit measurement of perceptual defense in a situation of tachistoscopic recognition, with a separate measure of the amount of response suppression also available.

In individual sessions, *S*s were given the task of identifying words flashed tachistoscopically on a screen. The particular tachistoscopic setting used was well below 50% recognition threshold, but it had previously been demonstrated (on a group of "calibrating" *S*s) to result in slightly above chance recognition. Each *S* was handed a list of eight words and was told by *E* that every word flashed on the screen would be some word from that list. Four of the words on the list were emotional and had been chosen on the basis of that *S*'s reactions to a word-association test given in an earlier individual session (details of this procedure, and of other aspects of the study, can be found in Mathews [7]). The other four words were neutral[3] and were paired with the emotional words for length, approximate structure, and relative frequency of occurrence in the Thorndike-Lorge General Word Count. During the tachistoscopic task, therefore, *S* had a total of eight possible alternative hypotheses about which word would be flashed, but he had no prior knowledge for predicting *which* word would be presented on any one particular flash. Each *S* was confronted with 32 flashes or presentations of a word.

Unknown to *S*, however, only *two* emotional

[1] Now with Educational Testing Service, Princeton, New Jersey.

[2] The completion of this paper was made possible by Wertheimer's participation in the Behavioral Science Conference at the University of New Mexico, Summer, 1957, under Project AF49(638)-33 sponsored by the Behavioral Science Division, Air Force Office of Scientific Research, ARDC, Paul Walter, Jr., principal investigator.

[3] The criterion for "neutrality" here was simply the mutual agreement of the authors that the word was probably lacking in emotional connotations for most people. Words such as Sit, Chair, North, Four, Day, and Many were used. As will be seen below, positive results were obtained in spite of the weakness of this technique (compared, e.g., with a technique in which a word association method is used with each individual *S* for the assignment of neutral words, but which creates the problem of how to match emotional and neutral words structurally).

and *two* paired neutral words were *actually* flashed; each of these four words was presented eight times in random order. The other four words were *never* flashed. Thus, this procedure permitted the abstraction of four categories for classifying *S*'s responses (two words being represented in each category):

Present (Words actually presented)	Absent (Words not presented)
a Emotional words	*c* Emotional words
b Neutral words paired with words in *a*	*d* Neutral words paired with words in *c*

Since people with hysteric tendencies might be considered likely to be repressors (2) and hence to show perceptual defense, this technique was used with 20 *S*s scoring high on the hysteria (*Hy*) scale of the MMPI—selected from a group of 271 students (7). A group of 20 *S*s scoring high on the psychasthenia (*Pt*) scale of the MMPI was also chosen for the experiment because it was considered that they might be hypersensitive to emotional material and, hence, might show perceptual vigilance (2).

The primary aim of this study was to obtain a single score for each individual which would be an index of "pure" perceptual defense—an index uncontaminated by any tendencies for response suppression. By correcting indices obtained from *a* and *b* by indices obtained from *c* and *d*, this aim could be accomplished. These indices were obtained as follows:

Through the two "Stimulus Absent" categories, *c* and *d*, an index of "pure" response tendencies is made available, since no stimulus corresponding to these four words was ever presented. One would expect that if *S* had no differential response tendencies, then 50% of his calls for words not actually presented would be calls of emotional words and 50% would be calls of neutral words.

This interpretation involves the crucial assumption of equal probability of calls for emotional and neutral words in the "Stimulus Absent" category. It might be argued that when *S* calls out a word that has *not actually been presented*, the nature of his error cannot be considered entirely independent of the stimulus which actually *was* presented. Thus the frequency of these two possible types of "misperceptions" (calls for emotional absent and neutral absent words) might be influenced by the relative similarity of the present stimuli to the absent stimuli. The argument is undoubtedly sound theoretically. However, a detailed analysis of calls actually made yielded the result that the variable of similarity between Present and Absent words had virtually no effect on our *S*s' "misperceptions." Aside from this analysis, there are further grounds for the assumption of an equal distribution of calls between Categories *c* and *d* if there is no response suppression. A priori, there is no reason to expect greater similarity of a presented word to an absent neutral than an absent emotional word; initial pairing of words in Category *c* with those in Category *d* had been made on the basis of similarity, so that a presented word would be about equally similar to the absent emotional and neutral words. Thus, in most instances at least, if a word actually presented to a given *S* (e.g., *suck*) resembled a word in Category *c* (e.g., *self*), it also resembled the paired word in Category *d* (e.g., *sale*). It can therefore be argued that any degree of undercalling (or overcalling) of emotional words as compared to neutral words can be interpreted as the degree of response suppression (or response sensitivity) manifested by *S*.

The expected number of calls of emotional words in Category *c* can therefore be calculated for each *S* by the formula $M = Np$, where N is the number of absent words called, and p is, by assumption, .5. Likewise, the standard deviation of total calls in the two "Stimulus Absent" categories can be calculated by the formula $\sigma = \sqrt{Npq}$.

In order to express each *S*'s deviation from a theoretically expected frequency of calls for emotional words in Category *c*, a "*z*" score $(X - M/\sigma)$ was used. The "*z*" score indicates both the extent and the nature of *S*'s response tendencies—a plus "*z*" score indicating response sensitivity, a zero "*z*" score indicating no differential tendencies, and a minus "*z*" score indicating response suppression.

However, in Categories *a* and *b* (words actually presented), relative calling of the two categories provides an index of *both* response tendencies *and* perceptual tendencies. The assumption that total "Stimulus Present" calls will be equally distributed between emotional and neutral categories (if neither response suppression nor perceptual defense occurs) still holds, since these words were in fact presented an equal number of times. Following the same procedure as described above, then, a "*z*" score can also be calculated for each *S* which represents his deviation from a theoretically expected frequency of calls for emotional words actually presented. Such a "*z*" score is an index that confounds both perceptual and response tendencies.

In order to obtain an index of "pure" perceptual tendencies, each *S*'s "*z*" score for "Stimulus Absent" words was subtracted from his "*z*" score for "Stimulus Present" words. Interpretation of the resulting difference "*z*" score for each individual can now be made on the basis of perceptual processes alone—since this is an index from which

TABLE 1

INDIVIDUAL "z" SCORES FOR *Hy* AND *Pt* SUBJECTS
(All scores are positive unless otherwise indicated)

Hy Group (Difference "z" scores)	*Pt* Group (Difference "z" scores)
1.92	−.48
−1.00	.99
−.50	−1.45
−.77	−.46
−.51	−1.88
.42	3.51
−.27	.80
−1.34	0.00
−.85	−.07
0.00	−.91
−.38	−1.84
.20	0.00
0.00	−.83
0.00	.80
−2.24	.09
0.00	0.00
0.00	−.27
−2.33	0.00
.25	−1.55
−.67	.28
$M = -.404$	$M = -.164$
CR = 1.94 ($p = .026$)	CR = .61 ($p = .271$)

the response tendencies of *S* have presumably been partialled out. Two assumptions are required to make this interpretation entirely tenable: namely, that response suppression (or response sensitivity) occurred to the same extent with the words actually presented as with those that were not presented,[4] and that response suppression effects and perceptual defense effects summate.[5]

In brief, the statistical procedure that was used in the study can be summarized as follows:

[Category *a*: Present Emotional "z" score (Perceptual + Response Tendencies)] − [Category *c*: Absent Emotional "z" score (Response Tendencies)] = [Difference Score (Pure Perceptual Tendencies)]

RESULTS

Tables 1 and 2 present the results in the form of distributions of the difference "z" scores for both *Hy* and *Pt* *S*s (Table 1) and in the form of group means and standard deviations (Table 2). The difference "z" scores are the result of algebraically subtracting the response suppression index from the contaminated index. As can be seen from the Tables, the *Hy* *S*s did manifest a perceptual defense effect over and above their response suppression effect. This effect was significant at better than the .03 level of significance.

TABLE 2

GROUP "z" SCORE MEANS AND SIGMAS

Group	Present Emotional "z"—from Categories *a* and *b*		Absent Emotional "z"—from Categories *c* and *d*		Difference "z"	
	M	σ	*M*	σ	*M*	σ
Hy ($N = 20$)	−.459	.96	−.056	.32	−.404	.91
Pt ($N = 20$)	−.386	.89	−.223	.73	−.164	1.17

As can also be seen from Tables 1 and 2, the *Pt* *S*s did not manifest the predicted perceptual vigilance (which would have been represented by a mean "z" score in the plus direction). Nor, however, did they show a significant perceptual defense effect, although there was a statistically insignificant trend in that direction. Since a *t* test showed the two groups did not differ from one another significantly (Diff. = .240, $t = .70$ $p > .10$), the results from both groups were combined, with the finding that twice as many *S*s exhibited perceptual defense (a negative "z" score) as exhibited perceptual vigilance (a positive "z" score).

SUMMARY

Using a tachistoscopic recognition technique, an experiment was designed that provided measures both of "perceptual defense" and of response suppression. The "perceptual defense" measure could be corrected by the response suppression measure to yield a "pure" perceptual defense measure. With this method, it was found that *S*s scoring high on the *Hy* scale of the MMPI demonstrate a significant perceptual defense effect over and above a response suppression effect, when confronted by words preselected so as to be emotional for each *S* individually. It is therefore concluded that perceptual defense effects cannot be accounted for entirely on the basis of response suppression.

REFERENCES

1. BLUM, G. S. Perceptual defense revisited. *J. abnorm. soc. Psychol.*, 1955, **51**, 24–29.
2. ERIKSEN, C. W., & DAVIDS, A. The meaning and clinical validity of the Taylor Anxiety Scale and the hysteric and psychasthenia scales from the MMPI. *J. abnorm. soc. Psychol.*, 1955, **50**, 135–137.
3. ERIKSEN, C. W., & BROWNE, C. T. An experimental and theoretical analysis of perceptual defense. *J. abnorm. soc. Psychol.*, 1956, **52**, 224–230.
4. ERIKSEN, C. W. Personality. In P. R. Farnsworth &

[4] This assumption seemed warranted in our study, since a counterbalanced technique was used to assign the four "most emotional" words for a given *S* to the "Stimulus Absent" and "Stimulus Present" category.

[5] Such a summation assumption may be oversimplified, but more complex comparable assumptions of an interactive kind would still yield results similar to those obtained.

Q. McNemar (Eds.), *Annual review of psychology.* Palo Alto, Calif.: Annual Reviews, 1957. Vol. 8.
5. GARNER, W. R., HAKE, H. W., & ERIKSEN, C. W. Operationism and the concept of perception. *Psychol. Rev.*, 1956, **63**, 149–159.
6. HOWES, D. H., & SOLOMON, R. L. A note on McGinnies' "Emotionality and perceptual defense." *Psychol. Rev.*, 1950, **57**, 229–234.
7. MATHEWS, ANNE. A "pure" measure of perceptual defense. Unpublished master's thesis, Univer. of Colorado, 1957.
8. POSTMAN, L. Perception, motivation, and behavior. *J. Pers.*, 1953, **22**, 17–31.
9. WHITTAKER, E. M., GILCHRIST, J. C., & FISCHER, J. W. Perceptual defense or response suppression? *J. abnorm. soc. Psychol.*, 1952, **47**, 732–733.

Received August 20, 1957.

PERCEIVED SIMILARITY AND THE PROJECTION OF SIMILARITY: THE INFLUENCE OF VALENCE[1]

ANTHONY J. SMITH
University of Kansas

OVER a period of years a system for conceptualizing interpersonal relations has been developed by Heider. It was first described in an article concerned with cognitive configurations and the conception of a balanced state (1). A more complete description of the system is now available (2).

The concept of structural balance can be stated in terms of a triadic relationship among a person "p" (who represents the subject), a second person "o," and an act, property, or characteristic "x." Though it is possible to extend the notion of structural balance to simpler or more complex situations, this paper is concerned only with the triad. Attention will be further confined to the triad in which one type of relationship exists among the variables: the relationship of liking or disliking, e.g., p L o (p likes o) or p∼Lo (p does not like o). The concept of balance may be illustrated very briefly as follows: When the subject p likes person o and o subscribes to opinion x, a force is generated such that p tends to agree with opinion x. If p agrees with x, the three elements p, o, and x, are in a state of balance. If p does not agree with statement x, an imbalanced state arises.

Horowitz, Lyons, and Perlmutter (3) have tested Heider's concept of balance in a field study of discussion groups. They found that *S*s are more acceptant of the acts of highly valued persons than they are of the acts of persons who are less valued. Further, they discovered that the *S* attributed to other accepted persons an acceptance of these same acts and were more likely to attribute to disliked persons a rejection of the acts. These results were consistent with the investigators' predictions.

Jordan (4) confronted *S*s with a series of relatively abstractly stated triadic relationships that were predicted to vary along a balance-imbalance continuum. Ratings of pleasantness were obtained from *S*s as indicators of degree of balance. In general, situations predicted to be balanced were found to be more pleasant than those predicted to be unbalanced, though some systematic exceptions were encountered.

In a later investigation, Smith (5) asked *S*s to examine partially completed test booklets (Revised Allport-Vernon) of hypothetical persons whose answers were either consistently similar or dissimilar to the *S*'s own. On the basis of the information provided, *S* completed the test booklets as he thought the hypothetical individuals would and later indicated his degree of acceptance of these individuals as work and leisure-time associates.

It was found that high valence (acceptance) was attributed to persons who were similar, whereas zero valence was attributed to those who were dissimilar or less similar. Moreover, greater additional similarity was projected on the more accepted persons than on the less accepted persons. It was reasoned on the basis of Heider's theory that this was a consequence of the differential in valences in the hypothetical o's.

In the study cited, the value scale employed to determine the extent of projection of similarity was that having the lowest valence for the *S*. The present investigation was designed in part to test the hypothesis that greater similarity is projected on similar persons than on dissimilar persons for both high and low valence materials.

A more important question concerns the differential influence of the valence of the test scale. It is hypothesized that the difference between the degree of similarity projected on a hypothetical similar individual and that projected on a hypothetical dissimilar individual is greater for

[1] This study was supported by a grant from the University of Kansas Research Fund. The author wishes to acknowledge the assistance of Harrison Madden and Ronald Sobol.

high valence material than for low valence material.

METHOD

Subjects

Ss in this study were 104 students enrolled in undergraduate courses in psychology and human relations. Forty-nine *Ss* were assigned to the low valence test scale group (including the 28 students reported in the original study [5]); the remaining 55 were assigned to the high valence test scale group. All *Ss* were required to perform the task twice under conditions of high and low similarity respectively. Forty-nine *Ss* in the high valence group were also required to respond to an intermediate degree of similarity.

Test Instrument

During the first two weeks in the semester, the Revised Allport-Vernon Scale of Values was administered as a routine part of the courses in general psychology, personality and cases in human relations from which the experimental *Ss* were drawn. The test booklets were scored by research assistants and the scales on which the students received the highest and lowest scores were determined.

The procedure by means of which hypothetical similar and dissimilar persons were created has been described in detail elsewhere (5). It involved preparing partially completed experimental test booklets based on the answers of the *S*. Information concerning the scale on which a particular *S* received the highest score (high valence test scale group) or the lowest score (low valence test scale group) was systematically eliminated from the experimental test booklet provided for him.

The information that was presented in the test booklet was designed either (*a*) to duplicate the answers of the *S*, thereby creating an hypothetical similar person or (*b*) to provide the maximum systematic difference from *S*'s answers, presumably yielding an hypothetical dissimilar person or (*c*) to yield a moderate systematic resemblance to the *S*'s answers resulting in a moderately similar hypothetical person.

Instructions

Six to ten weeks following the administration of the original test, *Ss* were met either individually or in small groups. They were reminded that in their everyday life they were always required to make judgments concerning characteristics of persons about which they had no direct knowledge on the basis of other information that they did possess. This experiment was designed to test their accuracy in this function. They would be provided with two or three partially completed test booklets filled out by other persons. Their task was to study each booklet individually until a clear impression of the original person emerged and then to complete the test booklet as the original person had.

ANALYSIS AND RESULTS

Similarity and the Projection of Similarity

The amount of similarity projected on each of the two or three hypothetical persons was defined as being inversely related to the absolute difference between the test score earned by the *S* on the critical scale and the score on the same scale in the experimental test booklet as completed by the *S*. This latter score was determined solely by *S*. Thus, the smaller the difference the greater the similarity attributed to the hypothetical person.

TABLE 1

MEANS OF DIFFERENCES IN AMOUNT OF SIMILARITY PROJECTED ON HYPOTHETICAL SIMILAR INDIVIDUALS AND HYPOTHETICAL DISSIMILAR INDIVIDUALS FOR TWO LEVELS OF TEST SCALE VALENCE

Valence	*N*	Mean Difference	σ_M	CR	*P*
Low Valence test scale	49	−4.78	1.28	3.73	.0002
High Valence test scale	55	−9.12	1.49	6.12	<.0001

The results of the analysis of low valence test scales are found in Table 1. The average difference between the similarity attributed to similar individuals and that attributed to dissimilar individuals is found to be −4.78 with a standard error of 1.28. In testing the significance of this difference from zero, a critical ratio of 3.73 is found with a corresponding probability of .0002.

An equivalent analysis for the high valence test scales is also reported in Table 1. The mean difference between amount of similarity attributed to similar individuals and that attributed to dissimilar individuals is −9.12, with a standard error of 1.49. The resulting critical ratio is 6.12, with a probability of <.0001.

A second form of analysis was also used in testing the first hypothesis. For 49 *Ss*, scores were available for three levels of similarity on the high valence test scale. If the hypothesis were to be supported, the absolute difference between own score and projected score should be lowest for similar test booklets, intermediate for the booklets showing moderate similarity and highest for the dissimilar test booklets. The three absolute difference scores for each person were ranked and the coefficient of concordance computed for the resulting sets of ranks. A coefficient of .41 was obtained. χ_r^2 was found to be 23.64, with a probability value of less than .001.

Differential Effects of Valences of the Test Scales

The second hypothesis was tested very simply by determining whether the mean difference for the high valence test scale (−9.12) was significantly greater than the mean difference for the low valence test scale (−4.78). The standard

error of the difference between these mean differences is found to be 1.96. The resulting critical ratio is 2.21 with a probability of <.03.

While the experiment was in progress, and during the discussions that followed the experimental sessions, questions concerning the test booklets were raised either publicly or privately by only ten *S*s. Thus it would seem that most *S*s remained naive throughout the experiment. These questions ranged from "Are these real people?" to "Was one of these mine?" On the assumption that these suspicions might have influenced the results, the data were reanalyzed excluding these ten *S*s. A critical ratio of 2.26 ($p < .05$) was obtained for the remaining 94 *S*s.

DISCUSSION

The theory of structural balance proposes that under the condition pLo, balance is attained when the remaining two elements of the triad are of equivalent valences, that is pLx, oLx, or p∼Lx, o∼Lx. Conversely, under either of the latter conditions, balance is achieved when pLo. This is precisely the situation that exists in the present experiment with respect to the hypothetical similar individual. When the *S* inspects the partially completed test booklet, each item establishes either the condition pLx, oLx or the condition p∼Lx, o∼Lx. Consequently, each item contributes an increment to the force in the direction of pLo.

When the *S* is obliged to complete the test booklet, the condition under which he works is pLo. Balance is attained if he attributes his own answers to the hypothetical person: pLx, oLx or p∼Lx, o∼Lx.

Moreover, the theory notes that when the circumstances are such that pLx, o∼Lx or vice versa, balance results if p∼Lo. Jordan has noted that all of his triadic relationships that have the component p∼Lo are rated as being quite unpleasant. It is conceivable that such a resolution of imbalance is less likely to occur than one that reduces the relation of p and o to the level of indifference. Under either circumstance o's valence for p is less for dissimilar persons than for similar persons. The trend toward attributing identical answers to the other person should be diminished as a consequence.

The results of the test of the first hypothesis conform to this logical analysis. Greater similarity is attributed to persons perceived as being similar than to persons perceived as being dissimilar. This result holds both when the attributed similarity and dissimilarity are on high valence and low valence characteristics.

The second hypothesis is concerned with the question of whether or not the relationship between degree of perceived similarity and extent of projected similarity is a function of the valence of the material to be projected. This situation can be translated into terms of Heider's system as follows: In a particular p, o, x triad, individual o and characteristic x have specific valences for p. The extent of the force generated such that p perceives either oLx or o∼Lx is a function of the valence of o (as demonstrated earlier), and the valence of x. The more crucial and acceptable (positive valence) x is for p, the more likely is p to associate or disassociate x and o. If o is not particularly crucial for p, the forces generated to establish balance are minimal.

The experimental results are consistent with this interpretation. Greater differentiation is made between similar and dissimilar persons on high valence items than on low valence items.

It should be noted that the employment of the Allport-Vernon scales restricts the treatment of results to relativistic terms. It is reasonable to believe that a person who received his highest score on the theoretical scale is positively oriented toward theoretical matters. However, a low score for the same person on the religious scale does not necessarily imply a rejection of religious issues and activities. It might involve rejection, indifference, or mild acceptance. The only legitimate statement concerning the theoretical and religious scales in this case would be that the former stands higher in positive valence than the latter.

SUMMARY

An attempt was made to investigate some of the implications of Heider's concept of balance.

The subjects were 104 students enrolled in undergraduate courses. On the basis of their answers to the items on the Revised Allport-Vernon Scale of Values, partially completed test booklets were prepared incorporating different degrees of similarity to each subject's original answers. The subjects were required to examine the test booklets and complete them in the way the hypothetical persons had.

The results were consistent with the hypotheses which stated that greater similarity is projected on similar persons than on dissimilar persons for both high and low valence materials and that the differential is significantly greater on high valence material than on low valence material.

REFERENCES

1. HEIDER, F. Attitudes and cognitive organization. *J. Psychol.*, 1946, **21**, 107–112.
2. HEIDER, F. *The psychology of interpersonal relations.* New York: Wiley, 1958.
3. HOROWITZ, M. W., LYONS, J., & PERLMUTTER,

H. V. Induction of forces in discussion groups. *Hum. Relat.*, 1951, **4**, 57–76.
4. JORDAN, N. Behavioral forces that are a function of attitudes and of cognitive organization. *Hum. Relat.*, 1953, **6**, 273–287.
5. SMITH, A. J. Similarity of values and its relation to acceptance and the projection of similarity. *J. Psychol.*, 1957, **43**, 251–260.

Received August 28, 1957.

FUTURE TIME PERSPECTIVE, OPTIMISM, AND ACADEMIC ACHIEVEMENT[1]

JOHN E. TEAHAN

Connecticut State Hospital

PREVIOUS studies have indicated that the high academic achiever seems to be characterized by more optimistic attitudes and a greater concern for future goals (2, 3, 4). The setting up of goals implies a temporal orientation that is geared to the future, since a person's goals imply expectations and anticipations of future successes. High academic achievers might therefore be expected to show an essentially future oriented approach to life, characterized by a time perspective with broader scope than that of the less successful student. Findings of this nature have already emerged from a study by Teahan and Drews (7) in which it was found that intellectually gifted and high achieving junior high school students had a more extensive total time perspective than low achieving students of similar intellectual levels. In the following study, however, only future time perspective was considered. It was hypothesized that high achievers would have a more predominant and extensive future time perspective than low achievers and that a positive relationship would be found between optimism and extensiveness of future time perspective.

METHOD

The subjects (*Ss*) used in this investigation were 60 seventh- and eighth-grade boys. Thirty *Ss* (high achievers) had received academic grades in the upper quartile of their classes for the previous two years; the remaining 30 *Ss* (low achievers) had been in the lower quartile for the same period. Groups were equated for age and for socioeconomic status as determined by rating the occupations of their fathers according to the method developed by Warner, Meeker, and Eells (8). Intellectual level was measured by the California Primary Mental Maturity Test which was administered in group form. Although no attempt was made to control for differences in intelligence between high achievers (mean IQ 108.3) and low achievers (mean IQ 95.5), the correlations obtained between intellectual level and the various measures of time perspective were found to approach zero in all cases (ranging from very low minus to very low plus correlations).

Three types of instruments were used to measure time perspective. The first of these techniques was originally developed by Eson (1) and consists of asking the *S* to record 25 things he had thought about or talked about in the past two weeks. *S* then rated these items according to whether, at the time he had thought or talked about them, they referred to something in the past, present, or future.

The second method, a story-completion technique, was first used by LeShan (5) for collecting data on time perspective. *S* was required to write a story starting with a partially completed statement or series of statements. The following sentences were used in the present investigation: (*a*) At three o'clock one bright sunny afternoon in May, two men were out walking near the edge of town . . . ; (*b*) Joe is having a cup of coffee in a restaurant. He's thinking of the time to come when . . . .

In the third part of the examination, *Ss* were presented with three TAT cards. These included Card 1 (boy with a violin), Card 12B (boy sitting alone before a cabin) and Card 14 (silhouette of a person in a window). The cards were presented one at a time in this order, and *Ss* were instructed to "Write a story about this picture."

When the entire testing session was ended, the examiner went over each story (both the story-completions and TATs) with the *S*, and asked the following questions: "How much time was involved in the action of this story—not in writing it but in the action described? How long would it have all taken if it had really happened?" In addition to this, each TAT story was rated by the examiner on a five-point scale according to the amount of optimism reflected in the outcome. Two other judges re-rated the TAT stories of 33 *Ss* (16 high and 17 low achievers) in a reliability check of ratings on this variable. The obtained correlations were .85, .91, and .91.

RESULTS

Reliability of Future Time Perspective Measures

Of all instruments used, the story-completion technique was least reliable in terms of correlations between scores based on different stories. When Story 1 was compared to Story 2, the obtained cor-

[1] This paper is based on a doctoral dissertation submitted to the department of psychology, Michigan State University, in 1957. It was also presented at the meeting of the APA in New York, August, 1957. The writer wishes to express his appreciation to his chairman, A. I. Rabin, and to the other members of the committee who aided in the completion of the study.

TABLE 1

COMPARISON OF HIGH AND LOW ACHIEVERS WITH RESPECT TO AMOUNT OF FUTURE TIME PERSPECTIVE

Instrument	High Achievers Medians	Low Achievers Medians	Z	P
Future References	13.5	9.2	2.77	.003
Story-Completion #1	4 hours	1.5 hours	2.03	.021
Story-Completion #2	3 days	6 hours	1.32	.093
All Story-Completions	3.5 days	1.5 days	1.69	.045
TAT Story #1	1 day	1.5 hours	1.82	.034
TAT Story #2	2.5 days	2 hours	2.02	.022
TAT Story #3	12 hours	1 hour	1.82	.034
All TAT Stori s	6 months	14 hours	2.21	.013

Note.—The Mann-Whitney nonparametric test was used because of the great variability in the time measures obtained which ranged from 5 minutes to 60 years. Because of the fact that the results were all in the predicted direction, and could not have been explained if they were different from that which was hypothesized, the probabilities listed are for a one-tailed test of statistical significance.

relation failed to reach a level of statistical significance. However, comparison of the TAT stories gave statistically significant correlations in all cases but one, when the scores obtained on any particular card by either high achievers, low achievers, or high and low achievers combined, were compared with their scores on any of the two remaining cards. There was also a high degree of internal consistency with respect to the future references obtained from recent conversations. An odd versus even comparison was made of the items listed by each *S* (only future references were considered), and the obtained correlations were significant well above the .01 level in all cases.

Future Time Perspective and Achievement

The recent conversations of high achievers included a significantly greater number of future references ($p < .01$) than those of low achievers. Although high achievers were found to have a more extensive future time perspective on both story-completions, only one of these (Story 1) gave statistically significant differences. When the results of both story-completions were combined, however, the differences were again significant by one-tailed test. The high achievers also demonstrated significantly more extensive future time perspective on all three stories given to the TAT cards.

Optimism and Future Time Perspective

A rank-difference correlation was computed between future time perspective and optimism on the TAT cards. The obtained coefficients of correlation were significant when high and low achievers were compared separately (coefficients of .47 and .40 respectively), or when they were studied as a total group (.50). Highly optimistic people tend to be more future oriented.

DISCUSSION

The results of Eson's (1) technique seems to indicate that high academic achievers are predominantly "anteverts" in so far as their recent thoughts and conversations are concerned, i.e., they tend to look mostly to the future. High achievers were also found to have a more extensive future time perspective than low achievers on most of the projective tests. These findings may mean that low achievers have a short-range or emergency type of future orientation, while high achievers have a long-range or foresightful type. Thus while continued and concentrated effort and attention to school subjects may be tedious and boring to the low achiever, these chores to the high achiever may represent meaningful subgoals on a time continuum in which the final end is eventual vocational success.

Students high in future extension also appeared to be more optimisitc. This agrees with Lewin's (6) theories concerning the relationship between self-confidence, hope, and a planning for a better future. The person without hope and confidence concentrates his attention on his immediate surrounds gaining whatever satisfactions he can in the present. In similar fashion, the student who has a history of poor grades might perceive the future as one in which only marginal success could be expected. Thus optimism, or the hope that positive changes can be expected in the future, would be intimately related to the extensiveness of a person's plans.

REFERENCES

1. ESON, M. E. An analysis of time perspective and five age levels. Unpublished doctoral dissertation, Univer. of Chicago, 1951.
2. GOUGH, H. G. Predicting success in graduate training. Inst. of Pers. Assessment and Res., Univer. of Calif., Berkeley, 1952. (Mimeo.)
3. GOUGH, H. G. Some theoretical problems in the construction of practical assessment devices for the early identification of high level talent. Dept. of Psychol. and Inst. of Pers. Assessment and Res., Univer. of Calif., Berkeley, 1953. (Mimeo.)
4. GOUGH, H. G. What determines the academic achievement of high school students. *J. educ. Res.*, 1953, **46**, 321–331.
5. LESHAN, L. L. Time orientation and social class. *J. abnorm. soc. Psychol.*, 1952, **47**, 589–592.
6. LEWIN, K. Time perspective and morale. In G. Watson (Ed.), *Civilian morale*. New York: Houghton Mifflin, 1942.
7. TEAHAN, J. E., & DREWS, ELIZABETH M. Time perspective and academic achievement. Paper read at Michigan Acad. of Arts and Sci., Detroit, March, 1957.
8. WARNER, W. L., MEEKER, MARCIA, & EELLS, K. *Social class in America*. Chicago: Science Research Assoc., 1949.

Received September 30, 1957.

HYSTERICS AND DYSTHYMICS AS CRITERION GROUPS IN THE MEASURE OF INTROVERSION–EXTRAVERSION: A REJOINDER TO EYSENCK'S REPLY[1]

JOHN J. SIGAL
Hampstead Child-Therapy Clinic, London

KOLMAN H. STAR
New South Wales University of Technology, Newcastle University College, N.S.W. Australia

AND CYRIL M. FRANKS
New Jersey Neuro-Psychiatric Institute

We hope to demonstrate that it was Eysenck (2), not we (3), who committed errors of logic and arrived at incorrect conclusions.

It was not Jung, but Eysenck who has contended that hysterics may be grouped with psychopaths and who coined the term dysthymia to include disorders not included under the term psychasthenia. Our findings, which Eysenck has accepted and supported by further (though inappropriate) data, demonstrated that hysterics may be grouped with dysthymics as justifiably as with pscyhopaths, but should probably not be grouped with either on I-E (introversion–extraversion). Contrary to Eysenck's statements, this finding is as relevant to his theory as it is to Jung's.

Eysenck seems to have made an error in logic. He has not recognized the difference between a necessary and a sufficient condition in logical argument. In order to demonstrate that hysterics are extraverted, it is a necessary condition that they be "toward the extraverted end of the scale" compared to dysthymics (assuming dysthymics to be introverted)—but it is not a sufficient condition. Elsewhere (1, p. 11 ff.) Eysenck, in quoting Jung, clearly demonstrated that Jung never intended to define an extravert as a less introverted introvert. To establish that hysterics are extraverted, it must also be shown that they are "toward the extraverted end of the scale" with respect to normals or *ambiverts* as Eysenck has called them, precisely as portrayed by him (1, p. 30).

From an experimental point of view, the relative positions of hysterics and dysthymics with respect to normals is important in determining the range of the I-E continuum over which one may consider as valid tests constructed, or conclusions about theories tested, using the former groups. Our results suggest that the use of these groups would limit the range to the introverted side of the continuum, that it would provide no information about the validity of tests or theories in the extraverted end of the continuum, and that tests constructed using hysterics and dysthymics as criterion groups might produce markedly skewed distributions when used in the normal population. These groups, used as a pair are, therefore, inadequate as criterion groups in the measure of I-E—if the E scale is a valid measure of I-E. The position of the normal group is essential to this conclusion. If we switched from one of Eysenck's deductions from Jung's theory to the other, it was because of logical and experimental necessity, not from want of logical consistency.

Since Eysenck has incorrectly anticipated our reply to his first point and has accepted the essential correctness of our results, his second point merits little discussion. Even allowing for the error in the logic of his first point, a significant difference between hysterics and psychopaths on I-E could hardly be considered to agree with prediction. It was because *our findings did not agree with prediction* that the hysteric–psychopath group (extraverted in theory but not in fact) was broken into its component parts, thereby producing smaller numbers. We might, indeed, have doubted the value of results based on so few cases were they not found to be in agreement with the results of other studies.

Eysenck suggested that we deliberately withheld relevant data, and that inclusion of these data would have altered our findings and the conclusions possible from these findings. We refer to his remarks about Jensen's data and the mean scores of the group used in standardizing the MPI.

Jensen selected his patients on the basis of extreme scores on I-E and obtained the diagnoses of these extreme scorers only. Our patients came from an entirely different hospital and were selected on the basis of diagnosis. The data from two such studies cannot be meaningfully combined. It was

[1] Editorial policy severely limits the space available to us. We are therefore only able to deal with some of the major issues raised by Eysenck, and even then in but a cursory manner.

Eysenck added his note of caution to ours (3) concerning the inadequacy of our sample for testing the merits of the N scale. Our summary did not repeat this note of caution contained in the text, so it is as well that Eysenck did. As the title of our article and the contents of Eysenck's reply indicated, however, the N scale results were neither the main nor an important issue in the study, so they will not be discussed in this rejoinder.

Eysenck who erred in pooling Jensen's data with those of Claridge, for example, which were obtained from unselected diagnosed patients. The reason for our use of Star's normal group rather than the standardization population was that the exact instructions and method of administering the test to the former group were known and could be followed, whereas there was no guarantee of the uniformity of these conditions in the testing of the latter group. In any case, the difference between the mean *E* scores of the two groups was too slight to warrant mentioning (12.63 compared to 12.45), since it would merely have confirmed a point for which we already had sufficient outside evidence.

In the paragraph before the summary, Eysenck indicated acceptance of our results, if not of our conclusions. If, as he suggested in this paragraph, there is a negative correlation between E and N for high N scorers and if, as he suggested, our diagnostic procedure resulted in the selection of less neurotic hysterics than are generally found in hospitals, then our selection was biased in favor of his theory, and presumably the result of a procedure which eliminated this bias would be even less favorable to his theory.

Finally, it should be pointed out that accepting the validity of the MPI as a measure of I-E would seem the least parsimonious of the alternatives open to Eysenck since, by so doing, he casts doubt on much of his earlier work in which hysterics were found to be extraverted.

SUMMARY

Eysenck has accepted our results despite their alleged premature publication, but has disputed our interpretation of them. The most prematurely published result, using his criterion of sample size, showed psychopaths to be extraverted, as predicted by his theory. Presumably, he did not intend this to be a contentious point. A consideration of the other arguments advanced by Eysenck, some of which were supported by irrelevant data, enables us to reaffirm our original statement, namely, either hysterics and dysthymics cannot be used as criterion groups in the measure of introversion–extraversion, or the E scale is not a good measure of introversion–extraversion, or both statements are true.

REFERENCES

1. EYSENCK, H. J. *The structure of human personality.* London: Methuen, 1953.
2. EYSENCK, H. J. Hysterics and dysthymics as criterion groups in the measure of introversion—extraversion: A reply. *J. abnorm. soc. Psychol.*, 1958, **57**, 250–252.
3. SIGAL, J. J., STAR, K. H., & FRANKS, C. M. Hysterics and dysthymics as criterion groups in the measure of introversion—extraversion. *J. abnorm. soc. Psychol.*, 1958, **57**, 143–148.

Received July 23, 1958.

PSYCHIATRY

JOURNAL FOR THE STUDY OF INTERPERSONAL PROCESSES

Contents of No. 3, Vol. 24, Aug., 1961

Kenneth Soddy: Isolation in the Mentally Retarded

Gerard J. Jones: Language of the Body

James Wiseman: Problems of Communication and Feelings in Psychotherapy

D. Wynne Goldsmith, Jean Macgregor: The Involvement of the Psychiatrist in Community Activities

Roy M. Waggoner and Dorothy Hall: Sex Role and Social Rank

Dwight E. Sargent: Family Environment and Schizophrenia

Nicholas D. Kiritsis and Thomas Ungvarski: Psychiatric Patients in Solitary Confinement

Book Notices

Books Received

Local Events

WILLIAM ALANSON WHITE PSYCHIATRIC FOUNDATION

1610 New Hampshire Avenue, N.W.

Washington 9, D.C.

PSYCHOPATHOLOGY

A Source Book

Edited by CHARLES F. REED, IRVING E. ALEXANDER, and SILVAN S. TOMKINS

Planned as a secondary text for courses in abnormal psychology and psychiatry, this selection of papers from recent literature in the field provides the student with an appreciation of the scope of present investigations. Most of the papers represent advances made within the last five years, and all have been selected on the basis of individual excellence, rather than for representativeness.

Illustrated. $12.50

Through your bookseller, or from

HARVARD UNIVERSITY PRESS

79 Garden Street, Cambridge 38, Mass.

AFTER-EFFECT TEST APPARATUS

for Testing Both Children and Adults

Complete equipment for administering the Spiral After-Effect Test. Recently published research indicates high reliability and validity in discriminating brain injuries from emotional states and normals.

- Apparatus is Portable and Compact
- Reversible Motor—No Wired Connections Necessary
- Single Switch for Positive and Negative Effects
- Built-in "Flash" and "Auditory Signal" Devices for Uninterrupted Performance
- Uses Regular Flashlight Batteries
- Stainless Rolls to Record Specifications

Complete with Spiral Instruction Manual, Subject Record Forms and One Year Warranty

$41.50 net

Order Direct from

PSYCHOLOGICAL RESEARCH & DEVELOPMENT CO.

[illegible] W. LAFAYETTE ST., [illegible] 3, FLORIDA